ACTA MATHEMATICA

ACADEMIAE SCIENTIARUM HUNGARICAE

ADIUVANTIBUS

REDIGIT

G. HAJÓS

TOMUS I. FASCICULUS 1

MAGYAR TUDOMÁNYOS AKADÉMIA
BUDAPEST, 1950

ACTA MATHEMATICA HUNGARICA

A MAGYAR TUDOMÁNYOS AKADÉMIA III. OSZTÁLYÁNAK MATEMATIKAI KÖZLEMÉNYEI

SZERKESZTŐSÉG ÉS KIADÓHIVATAL: BUDAPEST VI, SZTÁLIN-ÚT 31

Az Acta Mathematica orosz, francia, angol és német nyelven közöl értekezéseket a matematika köréből.

Az Acta Mathematica változó terjedelmű füzetekben jelenik meg: 20—30 ív terjedelemben, több füzet alkot egy kötetet. Évenként általában egy kötet jelenik meg.

A közlésre szánt kéziratok, lehetőleg géppel írva, a következő címre küldendők:

Acta Mathematica, Budapest 62, Postafiók 440.

Ugyanerre a címre küldendő minden szerkesztőségi és kiadóhivatali levelezés.

Az Acta Mathematica előfizetési ára kötetenként 60 forint. Megrendelhető a „Kultúra" Könyv és Hírlap Külkereskedelmi Vállalatnál (Budapest V, Akadémia-u. 10. Bankszámla: 929040. sz.), vagy külföldi képviseleteinél és bizományosainál.

„Acta Mathematica" издает трактаты из области математических наук на русском, французском, английском и немецком языках.

„Acta Mathematica" выходит в брошюрах переменного обьема (20—30 печатных листов) несколько выпусков объединяются в одном томе.

Ежегодно предвидено издание одного тома.

Предназначенные для публикации авторские рукописи следует направлять, по возможности машинописью, по следующему адресу:

Acta Mathematica, Budapest 62, Postafiók 440.

Поэтому же адресу направляется всякая корреспонденция для редакции и администрации.

Подписная цена „Acta Mathematica" — 60 форинтов за том. Заказы принимает Предприятие по внешней торговле книг и газет „Kultúra" (Budapest V, Akadémia-u. 10). Счет Банка № 929040) или его заграничные представительства и уполномоченные.

ВСТУПЛЕНИЕ

Обновление Венгерской Академии Наук открыло новую главу в истории венгерской науки. Ученые Венгрии всеми силами стремятся служить делу народа и своими исследованиями способствовать созидательному труду построения социализма. Венгерская Народная Республика оказывает развитию научной жизни нашей страны громадную материальную и моральную помощь и наука пользуется в нашей родине таким уважением и такой поддержкой, как еще никогда в нашей истории. Одной из характерных черт нашей обновленной науки является связь между научной теории и практической жизнью. Это взаимодействие оказывает серьезное, плодотворное влияние на развитие нашей науки.

Венгерская Академия Наук поставила себя целю изданием новой серии Acta Mathematica способствовать углублению международных связей прогрессивной науки, дальнейшему развитию науки, делу мира и прогресса и дружбы народов.

INTRODUCTION

La renaissance de l'Académie des Sciences de Hongrie ouvre un nouveau chapitre dans l'histoire des sciences hongroises. Les savants hongrois font tous leurs efforts pour servir la cause du peuple travailleur et aider par leurs travaux de recherche le travail créateur de l'édification du socialisme. La République Populaire Hongroise contribue largement, matériellement et moralement, au développement de la vie scientifique de notre pays. Dans notre pays, le travail scientifique jouit d'une estime et d'un soutien tels qu'il n'en a encore jamais joui au cours de notre histoire. Une des caractéristiques de notre vie scientifique renaissante est le contact entre la vie scientifique et la vie pratique de notre pays. Cette influence réciproque se fait fructueusement sentir dans le développement de notre vie scientifique.

Le but de l'Académie des Sciences de Hongrie, en publiant la nouvelle série des Acta Mathematica, est de contribuer par là au développement des relations internationales de la science progressiste, au développement de la science, à la défense de la Paix et du progrès, et au développement de l'amitié entre les peuples.

INTRODUCTION

The rebirth of the Hungarian Academy of Science has opened a new chapter in the history of Hungarian science. The scientists of Hungary endeavour in every way to serve the cause of the working people and with their research work to help in the creative task of building socialism. The Hungarian People's Republic affords vast help and encouragement to the development of the scientific life of our country and scientific work in Hungary today is honoured and aided to an extent that is unparalleled in the history of the land. One of the characteristic features of our reborne science is the connection between scientific theory and the practical life of the country. This interrelation has a profound stimulative effect on the development of our scientific life.

The aim of the Hungarian Academy of Science in starting the new series of Acta Mathematica is to contribute to the improvement of the international relations of progressive science, to the further development of science, to the ause of peace, progress and the closer friendship of the peoples.

EINLEITUNG

Die Wiedergeburt der ungarischen Akademie der Wissenschaften eröffnete einen neuen Abschnitt in der Geschichte der ungarischen Wissenschaft. Die ungarischen Gelehrten bemühen sich auf jede Art und Weise der Sache des werktätigen Volkes zu dienen und mit ihren Forschungen die schöpferische Arbeit des Aufbaues des Sozialismus zu fördern. Zur Entwicklung des wissenschaftlichen Lebens in unserem Lande trägt die ungarische Volksrepublik mit riesiger materieller und moralischer Hilfe bei. Die wissenschaftliche Arbeit in unserer Heimat wird in solchem Maße geschätzt und unterstützt, wie noch niemals in unserer Geschichte. Einer der charakteristischen Züge unserer wiedergeborenen Wissenschaft ist die Verbindung der wissenschaftlichen Theorie mit der Praxis im Leben unseres Landes. Diese Wechselwirkung ist von ernstem, fruchtbarem Einfluß auf die Entwicklung unseres wissenschaftlichen Lebens.

Mit der Ausgabe der neuen Serie der Acta Mathematica verfolgt die ungarische Akademie der Wissenschaften das Ziel, beizutragen zur Vertiefung der internationalen Verbindungen der fortschrittlichen Wissenschaften, zur Weiterentwicklung der Wissenschaft, zum Frieden und zum Fortschritt, zur Sache der engeren Freundschaft zwischen den Völkern.

ON THE FEUERBACH-SPHERES OF AN ORTHOCENTRIC SIMPLEX

By

E. EGERVÁRY (Budapest), member of the Academy

1. Attempting to extend the theorems of the geometry of the triangle to three and more dimensions it is known by experience that strict analogies exist only in the case of an orthocentric tetrahedron or simplex, i. e., such one whose altitudes have a point in common.

Actually it has been recognized rather long ago that the Feuerbach-circle (or nine-point circle) has no analogon in the case of a general tetrahedron, but if the tetrahedron is orthocentric then there are two spheres (the „twelve-point" spheres), each of which can be regarded as the three dimensional extension of the Feuerbach-circle.

Recent researches[1] have shown that with an orthocentric simplex in $n-1$ dimensions n spheres may be associated so that the k-th sphere contains the orthocenters and the barycenters of all the $k-1$ dimensional partial simplices. Consequently, in $n-1$ dimensions there are n extensions of the Feuerbach-circle. These Feuerbach-spheres belong to the pencil which is determined by the circumsphere and the polarsphere.

It is well known that the discussion of the Feuerbach-figure is rather asymmetrical and cumbersome if one uses Cartesian coordinates. Therefore several writers, when dealing with the Feuerbach-circle, made use of triangular coordinates. The discussion in Cartesian coordinates becomes even more clumsy in the case of three and more dimensions.

In the present paper we wish to show that the Feuerbach-figure in any dimension is capable of a symmetrical and very intuitive analytic representation by means of an orthocentric system of coordinates.

The orthocentric system of coordinates is a barycentric system whose fundamental simplex (tetrahedron) is orthocentric. In order to simplify the discussion of metric questions, the barycentric coordinates can be normalized in such a way that their sum should be equal to 1.

[1] E. Egerváry, On Orthocentric Simplexes, *Acta. Scient. Math. Szeged.*, 9 (1940), pp. 218—226.

These orthocentric coordinates proved to be a very useful instrument in the treatment of some metric questions[2]. Their application to the investigation of the Feuerbach-figure is particularly suggested by their close connection with the poly(penta)spherical coordinates.

2. The Feuerbach-circle possesses the following characteristic properties:

I. It is a Carnot-conic of the fundamental triangle belonging to the orthocenter and the barycenter.

II. It is a pedal-conic of the fundamental triangle belonging to the orthocenter and the circumcenter.

III. It touches each of the four tangent-circles of the fundamental triangle.

I shall show first that a simplex in $n - 1$ dimensions being given, one can associate with any pair of points n quadrics each of which is an extension of the Carnot-conic.

In agreement with this interpretation, the Feuerbach-spheres $\Phi_{(1)}$, $\Phi_{(2)}, \ldots, \Phi_{(n)}$ of an orthocentric simplex appear as the Carnot-quadrics belonging to the orthocenter and the barycenter.

Moreover I prove that the Feuerbach-sphere $\Phi_{(k)}$ is the pedal-quadric of the orthocenter and of its symmetric with respect to the center of $\Phi_{( )}$, i. e., it passes through the orthogonal projections of these points on the $k-1$ dimensional partial simplices.

Before entering into the discussion of the Feuerbach-figure we give a short summary of the properties of the orthocentric system of coordinates.

I. The orthocentric system of coordinates

1. If in the space of $n-1$ dimensions a fixed set of N ($\geqq n$) points $P_1, P_2, \ldots, P_N$ (containing at least one non-degenerate $n-1$ dimensional simplex) is given, then any point X may be represented as the barycenter of masses ξ_i placed in the points P_i ($i = 1, 2, \ldots, N$). Using the notation of GRASSMANN, a point X can be represented in the form

$$X = \frac{\xi_1 P_1 + \xi_2 P_2 + \ldots + \xi_N P_N}{\xi_1 + \xi_2 + \ldots + \xi_N}, \tag{1}$$

and the (real) numbers ξ_i are the homogeneous (and in the case $N > n$ *supernumerary*) barycentric coordinates of the point X. When dealing with metric problems it is convenient to use normal barycentric coordinates submitted to the restriction $\sum_1^N \xi_i = 1$ (with the exception of points at infinity).

[2] See e. g. E. EGERVÁRY, Über ein räumliches Analogen des Sehnenvierecks, *Journ. f. Math.*, 182 (1940) pp. 122—128.

The distance of two points $X = \Sigma \xi_i P_i$ and $Y = \Sigma \eta_i P_i$ is given by

$$\overline{XY}^2 = - \sum_{\substack{i,j=1 \\ (i>j)}}^{N} \overline{P_i P_j}^2 \, (\xi_i - \eta_i)\,(\xi_j - \eta_j), \qquad \sum_1^N \xi_i = \sum_1^N \eta_i = 1. \tag{2}$$

2. A general simplex of n points $P_1 P_2 \ldots P_n$ in $n-1$ dimensions is determined by $\frac{n\,(n-1)}{2}$ independent parameters, e. g. by the lengths of its edges $\overline{P_i P_j}$. If the simplex is orthocentric, i. e., if its altitudes meet in one point P_0 (the orthocenter), then its parameters have to satisfy $\frac{n\,(n-3)}{2}$ equations. In order to avoid the necessity of having regard continually to these equations of condition it seems to be desirable to determine an orthocentric simplex by a minimal number of independent and symmetrical parameters.

In one of my previous papers[1] I have shown that an orthocentric simplex $P_1 P_2 \ldots P_n$ in $n-1$ dimensions can be determined by n independent and symmetrical parameters in such a way, that the measure (length, volume) of its $k-1$ dimensional partial simplices ($k = 1, 2, \ldots, n-1$) is immediately given by the elementary symmetric functions of the parameters in the form

$$(k-1)!^2\, \overline{P_{j_1} P_{j_2} \ldots P_{j_k}}^2 = \lambda_{j_1} \lambda_{j_2} \ldots \lambda_{j_k} \left(\frac{1}{\lambda_{j_1}} + \frac{1}{\lambda_{j_2}} + \ldots + \frac{1}{\lambda_{j_k}}\right). \tag{3}$$

In particular

$$\overline{P_j P_k}^2 = \lambda_j + \lambda_k\,. \tag{3'}$$

From this representation it is obvious that any partial simplex of an orthocentric simplex is orthocentric too.

For some purposes it is more convenient to consider the set of $n+1$ points: the orthocenter P_0 and the vertices $P_1, P_2, \ldots, P_n$ as a whole, called an orthocentric set of $n+1$ points in the $n-1$ dimensional space, each point of the set being the orthocenter of the simplex formed by the others.

The mutual distances $\overline{P_i P_j}$ of the points of an orthocentric set can be expressed in the same way by $n+1$ symmetric parameters $\lambda_i \, (i = 0, 1, \ldots, n)$

$$\overline{P_i P_j}^2 = \lambda_i + \lambda_j \tag{3''}$$

but the n-dimensional measure of the simplex formed by the points $P_0 P_1 \ldots P_n$ of the $n-1$ dimensional space is equal to 0, consequently, the parameters λ_i are restricted by the relations

$$\frac{1}{\lambda_0} + \frac{1}{\lambda_1} + \ldots + \frac{1}{\lambda_n} = 0, \qquad (\lambda_i + \lambda_j > 0)\,. \tag{4}$$

According to this there is always one and only one negative value amongst the parameters λ_i, or, what is the same thing, one and only one

point of the set $P_0 P_1 \ldots P_n$ is contained in the interior of the convex cover of the set.

The consideration of the $n-1$ dimensional volumes of the simplices contained in the orthocentric set $P_0 P_1 \ldots P_n$ shows immediately that if the masses $\frac{1}{\lambda_i}$ are placed in the points P_i $(i = 0, 1, \ldots, n)$, then each point is the barycenter of the masses placed in the others. In other words, the mass-points $\frac{P_0}{\lambda_0}, \frac{P_1}{\lambda_1}, \ldots, \frac{P_n}{\lambda_n}$ satisfy the equations

$$\frac{P_0}{\lambda_0} + \frac{P_1}{\lambda_1} + \ldots + \frac{P_n}{\lambda_n} = 0\,, \qquad \frac{1}{\lambda_0} + \frac{1}{\lambda_1} + \ldots + \frac{1}{\lambda_n} = 0\,,$$

i. e. they constitute an indifferent mass-system.

3. In general we shall use in this paper the barycentric simplex-coordinates $x_1 x_2 \ldots x_n$ referring to the basic orthocentric simplex $P_1 P_2 \ldots P_n$. But, for the sake of simplicity and symmetry, at the beginning we shall develop the most important metric relations in terms of the supernumerary barycentric coordinates $\xi_0 \xi_1 \xi_2 \ldots \xi_n$ referring to the orthocentric set $P_0 P_1 P_2 \ldots P_n$. Afterwards we arrive by specialisation at the corresponding relations in terms of the simplex coordinates.

Let $\xi_0 \xi_1 \xi_2 \ldots \xi_n$ be a system of supernumerary barycentric coordinates of a point X, i. e., $X = \sum\limits_0^n P_i \xi_i \Big/ \sum\limits_0^n \xi_i$; $\sum\limits_0^n \xi_i \neq 0$. Then, from the fact that the mass-system P_i/λ_i is indifferent, we infer immediately that the most general system of supernumerary coordinates belonging to the same point X is given by

$$\xi_i' = \xi_i + \frac{\varrho}{\lambda_i}, \; (i = 0, 1, \ldots, n); \; X = \frac{\sum\limits_0^n P_i \xi_i}{\sum\limits_0^n \xi_i} = \frac{\sum\limits_0^n P_i \xi_i'}{\sum\limits_0^n \xi_i'} \tag{5}$$

where ϱ denotes an arbitrary real parameter.

Applying the expression (2) of the distance of two points and having regard to the relations (3″), we find immediately that the expression for the distance in terms of the coordinates ξ_i is given by

$$\begin{gathered} \overline{XY}^2 = - \sum \sum \overline{P_i P_j}^2 (\xi_i - \eta_i)(\xi_j - \eta_j) = \sum_0^n \lambda_i (\xi_i - \eta_i)^2 ; \\ \sum_0^n \xi_i = \sum_0^n \eta_i = 1 \,. \end{gathered} \tag{6}$$

I should not miss here to draw the attention to the analogy between the orthogonal Cartesian coordinates and the orthocentric coordinates. For

both system of coordinates (and only in these cases) the expression for the squared distance reduces to a sum of squares.

4. The equation of any sphere, having the center $C = \sum_0^n \gamma_i P_i$; $\sum_0^n \gamma_i = 1$ and the radius $r^2 = \sum_0^n \lambda_i \gamma_i^2$, may be derived immediately from (6)

$$\sum_0^n \lambda_i (\xi_i - \gamma_i)^2 - r^2 \equiv \sum_0^n \lambda_i \xi_i^2 - 2 \sum_0^n \lambda_i \gamma_i \xi_i = 0; \quad \sum_0^n \xi_i = \sum_0^n \gamma_i = 1,$$

or in homogeneous form

$$\sum_0^n \lambda_i \xi_i^2 - 2 \sum_0^n \lambda_i \gamma_i \xi_i \sum_0^n \xi_i = 0. \tag{7}$$

We infer from this that the $n+1$ spheres Γ_k having their centers in the vertices $P_k = \sum \delta_{ki} P_i$ ($\delta_{ki} = 0$ for $k \neq i$, $\delta_{kk} = 1$) and their radii $r_k^2 = \lambda_k$ are represented by the equations

$$\Gamma_k \equiv \sum_0^n \lambda_i \xi_i^2 - 2 \lambda_k \xi_k \sum_0^n \xi_i = 0, \qquad (k = 0, 1, 2, \ldots, n). \tag{8}$$

Any two of these spheres are orthogonal to each other in consequence of the equations $r_k^2 + r_h^2 - \overline{P_k P_h}^2 = \lambda_k + \lambda_h - (\lambda_k + \lambda_h) = 0$.

Comparing (7) with (8) we see that the equation of any sphere (point, plane) may be written as a homogeneous, linear combination of the forms Γ_k

$$\sum_0^n \gamma_i \sum_0^n \lambda_i \xi_i^2 - 2 \sum_0^n \lambda_i \gamma_i \xi_i \sum_0^n \xi_i \equiv \sum_0^n \gamma_i \Gamma_i = 0 \tag{9}$$

and in the case of this representation of a sphere the *supernumerary* coordinates of its center are proportional to the coefficients γ_i, while the radius is given by $r^2 = \sum_0^n \lambda_i \gamma_i^2 / (\sum_0^n \gamma_i)^2$.

The plane resp. the point are obviously characterised by the relations $\sum_0^n \gamma_i = 0$ resp. $\sum_0^n \lambda_i \gamma_i^2 = 0$.

5. In order to pass from the supernumerary coordinates $(\xi_0 \xi_1 \xi_2 \ldots \xi_n)$ to the simplex-coordinates $(x_1 x_2 \ldots x_n)$ referring to one of the simplices, e. g. to $P_1 P_2 \ldots P_n$, we have only to impose on the supernumerary coordinates the restriction that the point P_0 is deprived of mass. This involves that the arbitrary parameter ϱ in the expression (5) of the supernumerary coordinates must be chosen so that $x_0 = \xi_0 - \varrho/\lambda_0 = 0$, i. e. $\varrho = \xi_0 \lambda_0$. Hence the passage from ξ_i to x_j is mediated by the equations (and similarly the passage from γ_i to c_j)

$$x_j = \xi_j - \frac{\lambda_0}{\lambda_j} \xi_0 \quad (j = 1, 2, \ldots, n). \tag{10}$$

By means of these equations $\sum_0^n \lambda_i (\xi_i - \gamma_i)^2$ will be transformed into $\sum_1^n \lambda_j (x_j - c_j)^2$ and similarly $\sum_0^n \xi_i$ into $\sum_0^n x_j$; consequently, the general equation of a sphere in orthocentric simplex-coordinates takes the homogeneous form

$$\sum_1^n \lambda_j x_j^2 - 2 \sum_1^n a_j x_j \sum_1^n x_j = 0 \tag{11}$$

with arbitrary coefficients a_j.

Comparing this with (9), excluding the case of a plane and assuming $\sum_0^n \gamma_i = 1$, we get for the *supernumerary* coordinates γ_i and the radius r of the sphere the following expressions:

$$\gamma_0 = 1 - \sum_1^n \frac{a_j}{\lambda_j}; \quad \gamma_j = \frac{a_j}{\lambda_j}; \quad r^2 = \sum_1^n \lambda_i \gamma_i^2 = \lambda_0 \left(1 - \sum_1^n \frac{a_j}{\lambda_j}\right)^2 + \sum_1^n \frac{a_j^2}{\lambda_j}. \tag{12}$$

The corresponding formulae in terms of the orthocentric simplex-coordinates x_j follow from here immediately by substituting these values in the equations (10).

II. Representation of the Feuerbach-spheres as Carnot-quadrics

1. Two points

$$U = \frac{u_1 P_1 + u_2 P_2 + u_3 P_3}{u_1 + u_2 + u_3} \quad \text{and} \quad V = \frac{v_1 P_1 + v_2 P_2 + v_3 P_3}{v_1 + v_2 + v_3}$$

in the plane of the basic-triangle $P_1 P_2 P_3$ should be projected from the vertices on the opposite sides. According to a theorem of Carnot, these projections $(0\, u_2 u_3)$ $(u_1\, 0\, u_3)$ $(u_1 u_2\, 0)$; $(0\, v_2 v_3)$ $(v_1\, 0\, v_3)$ $(v_1 v_2\, 0)$ lie on a conic represented by the equation

$$2\left(\frac{x_1^2}{u_1 v_1} + \frac{x_2^2}{u_2 v_2} + \frac{x_3^2}{u_3 v_3}\right) - \left(\frac{x_1}{u_1} + \frac{x_2}{u_2} + \frac{x_3}{u_3}\right)\left(\frac{x_1}{v_1} + \frac{x_2}{v_2} + \frac{x_3}{v_3}\right) = 0.$$

Consider now two points U, V in the $n-1$ dimensional space, whose barycentric coordinates are $(u_1\, u_2 \ldots u_n)$ resp. $(v_1\, v_2 \ldots v_n)$.

These points U, V can now be projected from any $n-k-1$ dimensional partial simplex $P_{k+1} P_{k+2} \ldots P_n$ on the $k-1$ dimensional complementary partial simplex $P_1 P_2 \ldots P_k$. The projection of U (being the point of intersection of the $n-k$ dimensional plane $P_{k+1} P_{k+2} \ldots P_n U$ with the $k-1$ dimensional plane $P_1 P_2 \ldots P_k$) has the coordinates

$$(u_1\, u_2 \ldots u_{k-1}\, u_k\, 0\, 0 \ldots 0)$$

and the coordinates of the projection of V are

$$(v_1\, v_2\, \ldots\, v_{k-1}\, v_k\, 0\, 0\, \ldots\, 0).$$

Obviously all these projections of the pair of points U, V on the $k-1$ dimensional partial simplices lie on a quadric whose equation is easily found to be

$$k \sum_1^n \frac{x_j^2}{u_j\, v_j} - \sum_1^n \frac{x_j}{u_j} \sum_1^n \frac{x_j}{v_j} = 0. \tag{13}$$

Consequently, with any pair of points and with an $n-1$ dimensional (general) simplex n quadrics may be associated each of which can be interpreted as an extension of the Carnot-conic, i. e., the quadrics given by (13) for the values $k = 1, 2, \ldots, n$.

2. Let us now assume that the basic simplex $P_1 P_2 \ldots P_n$ is an orthocentric one, specified by the parameters $\lambda_1, \lambda_2, \ldots, \lambda_n$ and consider the Carnot-quadrics $\Phi_{(1)}, \Phi_{(2)}, \ldots, \Phi_{(n)}$ associated with the orthocenter $O = P_0 = \sum_1^n \frac{P_j}{\lambda_j} \Big/ \sum_1^n \frac{1}{\lambda_j}$ and the barycenter $B = \frac{1}{n} \sum_1^n P_j$.

Their equation follows at once from (13):

$$\Phi_{(k)} \equiv k \sum_1^n \lambda_j\, x_j^2 - \sum_1^n \lambda_j\, x_j \sum_1^n x_j = 0 \quad (k = 1, 2, \ldots, n)\,, \tag{14}$$

and comparing this equation with (11) we notice that they represent n spheres belonging to the pencil which is determined by the polarsphere ($k = \infty$) and the circumsphere ($k = 1$). But the projections of O and B are obviously identical with the orthocenters and barycenters of the partial simplices, consequently, each of the spheres $\Phi_{(1)}, \Phi_{(2)}, \ldots \Phi_{(n)}$ can be interpreted as an $n-1$ dimensional extension of the Feuerbach-circle. The sphere $\Phi_{(k)}$ ($k = 2, 3, \ldots, n-1$) passes through the orthocenters and barycenters of all the $k-1$ dimensional partial simplices, $\Phi_{(1)}$ is the circumsphere; $\Phi_{(n)}$ is the orthocentroidal sphere, because it has the join of the orthocenter O and barycenter B as diameter.

Applying (12) we get for the radius $r_{(k)}$ and the *supernumerary* coordinates $\gamma_i^{(k)}$ of the center $C_{(k)}$ of the Feuerbach-sphere $\Phi_{(k)}$

$$\begin{gathered} \gamma_0^{(k)} = 1 - \frac{n}{2\,k}; \quad \gamma_j^{(k)} = \frac{1}{2\,k} \quad (j = 1, 2, \ldots, n)\,; \\ r_{(k)} = \sqrt{(n - 2\,k)^2\, \lambda_0 + \lambda_1 + \ldots \lambda_n}/2\,k. \end{gathered} \tag{15}$$

According to (15) the center $C_{(k)}$ of $\Phi_{(k)}$ is given by

$$C_{(k)} = \frac{(2\,k - n)\, O + n\, B}{2\,k} = O + \frac{n}{2\,k}(B - O)\,, \tag{16}$$

i. e., $C_{(k)}$ is collinear with O and B and coincides with the terminal-point of the

vector $\dfrac{n}{2k}\overrightarrow{OB}$ issuing from O. Thus all the centers lie on the „Euler-line" joining O and B.

3. Having regard to the following discussions it is convenient to distribute the Feuerbach-spheres into complementary pairs $\Phi_{(k)}$, $\Phi_{(n-k)}$ (which coincide only for $n = 2m$, $k = m$). The radii of a complementary pair are connected by the simple relation

$$k\, r_{(k)} = (n-k)\, r_{(n-k)},$$

while their centers satisfy the equations

$$2kC_{(k)} = nB + (2k-n)O; \quad 2(n-k)C_{(n-k)} = nB - (2k-n)O,$$

i. e., the barycenter B, the orthocenter O and the centers $C_{(k)}$, $C_{(n-k)}$ of a pair of complementary Feuerbach-spheres form a harmonic quadruple of points.

Particularly in the case of the orthocentric tetrahedron ($n = 4$) we get from (15)

$$C_{(1)} = 2B - O;\ r_{(1)} = \sqrt{4\lambda_0 + \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4}/2 \quad \text{(circumsphere)}$$

$$C_{(2)} = B \quad ,\ r_{(2)} = \sqrt{\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4}/4 \quad \text{(first twelve-point-sphere)}$$

$$3\,C_{(3)} = 2B + O,\ r_{(3)} = \sqrt{4\lambda_0 + \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4}/6 \quad \text{(second twelve-point-sphere)}$$

in agreement with the well known results in the geometry of the tetrahedron.

It has been observed that the four circumcircles, resp. the four Feuerbach-circles of triangles contained in a planar orthocentric quadruple of points are equal, while in the three dimensional space no such relations exist. The expression (15) of $r_{(k)}$ shows clearly that similar relations exist only in a space of an even number of dimensions. In the case of $2m$ dimensions we have from (15)

$$r_{(k)} = \sqrt{\lambda_0 + \lambda_1 + \ldots + \lambda_{2m+1}}/2k \qquad (k = m,\ m+1),$$

i. e., the Feuerbach-spheres $\Phi_{(m)}$ resp. $\Phi_{(m+1)}$ of the orthocentric simplices contained in an orthocentric set of $2m+1$ points in $2m$ dimensions are equal.

III. Representation of the Feuerbach-spheres as pedal quadrics

1. An other characteristic feature of the Feuerbach-circle is that it passes through the orthogonal projections of the orthocenter and circumcenter on the sides of the basic triangle (it is the pedal-circle of O and $C_{(1)}$).

In order to establish the corresponding property of the Feuerbach-spheres in $n-1$ dimensions let us consider the intersection of the Feuerbach-sphere

$$\Phi_{(k)} \equiv k \sum_1^n \lambda_j x_j^2 - \sum_1^n \lambda_j x_j \sum_1^n x_j = 0$$

with one of „its" partial-simplices, e. g. with

$$x_{k+1} = x_{k+2} = \ldots = x_n = 0.$$

This intersection

$$\Phi^* \equiv k \sum_1^k \lambda_j x_j^2 - \sum_1^k \lambda_j x_j \sum_1^k x_j = 0$$

s obviously the orthocentroidal-sphere of the partial-simplex $P_1 P_2 \ldots P_k$. According to one of our former results **(II. 2.)** the center C^* of Φ^* is collinear with and equidistant to the orthocenter O^* and the barycenter B^* of the partial simplex $P_1 P_2 \ldots P_k$.

Erect now through O^*, C^* and B^* $n-k-1$ dimensional planes which are perpendicular to the $k-1$ dimensional plane of $P_1 \ldots P_k$. These parallel planes meet the Euler-line of the basic orthocentric simplex in O, $C_{(k)}$ and in a point $B_{(k)}$, this latter being obviously the symmetric of O with respect to $C_{(k)}$.

Starting with any other $k-1$ dimensional partial-simplex, we arrive evidently at the same point $B_{(k)}$. Hence we have the theorem:

The Feuerbach-sphere $\Phi_{(k)}$ is the pedal-sphere of the orthocenter O and of its symmetric $B_{(k)}$ with respect to the center $C_{(k)}$ of $\Phi_{(k)}$. In other words, $\Phi_{(k)}$ passes through the orthogonal projections of O and $B_{(k)}$ on all the $k-1$ dimensional partial-simplices.

The point $B_{(k)}$, being the symmetric of O with respect to $C_{(k)}$, is represented by

$$B_{(k)} = 2\,C_{(k)} - O = O + \frac{n}{k}\,(B - O),$$

i. e., $B_{(k)}$ coincides with the terminal-point of the vector $\frac{n}{k}\overrightarrow{OB}$ issuing from O.

All these centers of projection lie on the Euler-line, moreover $C_{(2k)}$ coincides with $B_{(k)}$.

IV. On the intersections of the Feuerbach-spheres and the altitudes

1. In the case of three and more dimensions the notion of the altitude may be viewed from a more general point of view, i. e., as the common perpendicular of a pair of complementary partial-simplices. This extension is justified by the fact that the normal-transversal of any pair of complementary partial-simplices passes through the orthocenter O as well as through the orthocenters $O_{(k)}$ and $O_{(n-k)}$ of the complementary partial-simplices.

Indeed, the orthocenter of $P_1 P_2 \ldots P_n$ is $O = \sum P_j/\lambda_j : \sum_1^n 1/\lambda_j$, the orthocenters of $P_1 P_2 \ldots P_k$ resp. $P_{k+1} P_{k+2} \ldots P_n$ are $O_{(k)} = \sum_1^k P_j/\lambda_j : \sum_1^k 1/\lambda$

resp. $O_{(n-k)} = \sum\limits_{k+1}^{n} P_j/\lambda_j : \sum\limits_{k+1}^{n} 1/\lambda_j$. Hence

$$O_{(k)} \sum_{1}^{k} 1/\lambda_j + O_{(n-k)} \sum_{k+1}^{n} 1/\lambda_j = O \sum_{1}^{n} 1/\lambda_j ,$$

i. e., O, $O_{(k)}$, $O_{(n-k)}$ are collinear.

We have further $\overline{OO_{(k)}} \perp \overline{P_1P_2\ldots P_k}$ and $\overline{OO_{(n-k)}} \perp \overline{P_{k+1}\ldots P_n}$, consequently the altitude $\overline{O_{(k)}\, O_{(n+k)}}$ belonging to $P_1 P_2 \ldots P_k$ and $P_{k+1} P_{k+2} \ldots P_n$ is the normal-transversal of these partial-simplices.

In this wider sense the number of the altitudes of an orthocentric simplex $P_1 P_2 \ldots P_n$ in the $n-1$ dimensional space is obviously

$$2^{2m} - 1 \text{ for } n = 2m+1 \text{ and } 2^{2m-1} + \frac{1}{2}\binom{2m}{m} - 1 \text{ for } n = 2m.$$

We shall prove that the products of the segments, into which any of these altitudes is divided by the orthocenter, has the same value, this common value being equal to $|\lambda_0|$.

The segment $\overline{OO_{(k)}}$ is the altitude (in the narrower sense) of the partial-simplex $P_0 P_1 \ldots P_k$, therefore its length is given by

$$\overline{OO}^2_{(k)} = k^2 \frac{\overline{P_0 P_1 P_2 \ldots P_k}^2}{\overline{P_1 P_2 \ldots P_k}^2} \text{ and similarly } \overline{OO}^2_{(n-k)} = (n-k)^2 \frac{\overline{P_0 P_{k+1} \ldots P_n}^2}{\overline{P_{k+1} \ldots P_n}^2}$$

Applying (3) we get from here

$$\overline{OO}^2_{(k)} \cdot \overline{OO}^2_{(n-k)} =$$

$$= \frac{\lambda_0 \lambda_1 \ldots \lambda_k \left(\dfrac{1}{\lambda_0} + \dfrac{1}{\lambda_1} + \ldots + \dfrac{1}{\lambda_k}\right)}{\lambda_1 \ldots \lambda_k \left(\dfrac{1}{\lambda_1} + \ldots + \dfrac{1}{\lambda_k}\right)} \cdot \frac{\lambda_0 \lambda_{k+1} \ldots \lambda_n \left(\dfrac{1}{\lambda_0} + \dfrac{1}{\lambda_{k+1}} + \ldots + \dfrac{1}{\lambda_n}\right)}{\lambda_{k+1} \ldots \lambda_n \left(\dfrac{1}{\lambda_{k+1}} + \ldots + \dfrac{1}{\lambda_n}\right)}$$

or, having regard to $\sum\limits_{0}^{n} 1/\lambda_i = 0$,

$$\overline{OO}^2_{(k)} \cdot \overline{OO}^2_{(n-k)} = \left(\lambda_0 + \frac{1}{\dfrac{1}{\lambda_1} + \ldots + \dfrac{1}{\lambda_k}}\right)\left(\lambda_0 + \frac{1}{\dfrac{1}{\lambda_{k+1}} + \ldots + \dfrac{1}{\lambda_n}}\right) = \lambda_0^2 ,$$

q. e. d.

2. The denomination „nine-point circle" is justified by the fact, that this circle passes not only through the projections of O and B on the sides, but it passes also through the middle-point of the „upper" segments of the

altitudes. A corresponding theorem is known in the case of an orthocentric tetrahedron, i. e., the sphere which contains the orthocenters and barycenters of the faces, divides the „upper" segments of the altitudes in the ratio 1 : 2.

These statements can be generalised as follows. Each extremity of an altitude is lying on a Feuerbach-sphere. E. g. the extremities of an altitude of type $\overline{O_{(k)} O_{(n-k)}}$ (belonging to a $k-1$ dimensional and to an $n-k-1$ dimensional partial-simplex) are lying on $\Phi_{(k)}$ and $\Phi_{(n-k)}$. Using our former results we can easily determine the second point of intersection of a Feuerbach-sphere with its corresponding altitude.

Retaining the previously used notations and denoting the second point of intersection of $\Phi_{(k)}$ with its altitude $\overline{O_{(k)} O_{(n-k)}}$ by $O'_{(k)}$, we inquire after the value of the ratio

$$\frac{\overline{OO'_{(k)}}}{\overline{OO_{(n-k)}}} = \frac{\overline{OO'_{(k)}} \cdot \overline{OO_{(k)}}}{\overline{OO_{(n-k)}} \cdot \overline{OO_{(k)}}}.$$

It has been previously proved that $\overline{OO_{(k)}} \cdot \overline{OO_{(n-k)}} = \lambda_0$, thus we have only to calculate $\overline{OO'_{(k)}} \cdot \overline{OO_{(k)}}$, i. e., the power of the orthocenter O with respect to the Feuerbach-sphere $\Phi_{(k)}$. Using our former results (15) we have

$$\overline{OO'_{(k)}} \cdot \overline{OO_{(k)}} = OC^2_{(k)} - r^2_{(k)} = \frac{n^2}{k^2} r^2_{(n)} - r^2_{(k)} = \lambda_0 \frac{n-k}{k}.$$

Hence we infer that $\overline{OO'_{(k)}} : \overline{OO_{(n-k)}} = (n-k) : k$ and this result can be stated as follows.

Each altitude of the type $\overline{O_{(k)} O_{(n-k)}}$ joins a point $O_{(k)}$ of the Feuerbach-sphere $\Phi_{(k)}$ to a point $O_{(n-k)}$ of the complementary Feuerbach-sphere $\Phi_{(n-k)}$. Consequently, this altitude cuts both of $\Phi_{(k)}$, $\Phi_{(n-k)}$ once more, and the positions of these intersections $O'_{(k)}$, $O'_{(n-k)}$ are determined by the ratios

$$\frac{\overline{OO'_{(k)}}}{\overline{OO_{(n-k)}}} = \frac{n-k}{k}; \quad \frac{\overline{OO'_{(n-k)}}}{\overline{OO_{(k)}}} = \frac{k}{n-k}.$$

This theorem contains the above mentioned results in the geometry of the triangle, resp. of the tetrahedron, if we adopt the circumcircle, resp. sphere as the first member $\Phi_{(1)}$ in the Feuerbach-sequence.

If $n = 2m$ and $k = m$, then the two complementary Feuerbach-spheres coincide. This is realized in three dimensions ($n = 4$) for $k = 2$, i. e., in the case of the first twelve-point sphere.

(Received 8 May 1950)

ШАРЫ ФЕЙЕРБАХА ОРТОЦЕНТРИЧНОГО СИМПЛЕКСА

Е. ЭГЕРВАРИ (Будапешт)

(Резюме)

Если ребра симплекса $P_1 P_2 \ldots P_n$ размерности $n-1$ могут быть выражены с помощью n параметров в виде $\overline{P_i P_j}^2 = \lambda_i + \lambda_j$, то симплекс является ортоцентричным и в относящейся к нему барицентричной системе координат $(x_1, x_2, \ldots, x_n)$ координаты ортоцентра суть $\left(\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \ldots, \frac{1}{\lambda_n}\right)$ и общее уравнение шара будет

$$a_0 \sum_1^n \lambda_i x_i^2 - \sum_1^n a_i x_i \sum_1^n x_i = 0.$$

Уравнение

$$k \sum_1^n \lambda_i x_i^2 - \sum_1^n \lambda_i x_i \sum_1^n x_i = 0$$

получающееся из соответствующей специализации коэффициентов $a_0 a_1 \ldots a_n$ означает такой шар, который содержит ортоцентры и барицентры всех частных симплексов $P_{r_1} P_{r_2} \ldots P_{r_k}$ всех размерностей $k-1$, то есть все шары соответствующие $k = 1, 2, \ldots, n$ могут считаться обобщением окружности Фейербаха. Использованием вышеприведенного уравнения шаров Фейербаха многие теоремы тригонометрии и геометрии тетраэдров обобщаются на ортоцентричные симплексы.

THÉORIE MATHÉMATIQUE DU TRAFIC DE MARCHANDISES SOUS LE RÉGIME DU CAPITALISME DE MONOPOLE

Par

GEORGES ALEXITS (Budapest), membre de l'Académie

1. Méthode de la fondation mathématique de l'économie politique

Le traitement des problèmes de l'économie politique par des méthodes mathématiques a déja un long passé. C'est bien compréhensible que les succès de la physique mathématique aient éveillé le désir de décrire d'une manière exacte les lois de l'économie politique. Malgré cela, les travaux de mathématiques économiques sont loin d'approcher l'exactitude de la physique théorique. La vérité est que la plupart de ces investigations ne cherche qu'à se donner l'apparence d'une recherche scientifique en se servant du symbolisme mathématique, mais en réalité elles ne font que d'habiller les allégations des différentes théories bourgeoises d'économie dans un vêtement mathématique. Mais même l'usage qu'elles font des formes mathématiques ne paraît point exact parce qu'il faut considérer en général les refontes mathématiques des notions de l'économie politique bourgeoise comme manquées en principe.

Avant tout, nous considérons *intenable* le point de vue des travaux bourgeois de l'économie mathématique qui examine les phénomènes économiques d'une façon axiomatique et par là ne tient pas compte de la structure sociale des phénomènes économiques qu'ils veulent décrire. De cette manière, les méthodes mathématiques employées dans l'économie bourgeoise donnent au système économique capitaliste l'apparence d'être la seule forme possible de l'économie. Le système économique capitaliste apparaît donc dans ces théories mathématiques de la science économique bourgeoise comme une nécéssité mathématique alors qu'il est évident que c'est entré dans la théorie non comme *conséquence* du contenu mathématique, mais parce que ce système a été tacitement *insinué* dans les formes mathématiques. Il est clair qu'une théorie qui croit trouver les mêmes lois économiques dans le système économique d'une société barbare dont la production est basée sur le travail des esclaves que dans une société développée du capitalisme financier est nécéssairement fausse et

ne pourrait être considerée comme science sérieuse malgré *l'apparence* mathématique.

En principe doit encore être considéré intenable un point de vue qui promet de traiter exactement les relations quantitatives se présentant dans des formes différentes de la valeur d'économie politique sans mettre au point la notion de valeur comprise implicitement dans les symboles mathématiques. Pourtant, nous rencontrons dans la science économique mathématique bourgeoise à tous les pas cette erreur: les formules expriment des changements de valeur, mais on ne peut pas établir qu'est-ce que c'est qui change: est-ce seulement la valeur des fonctions figurant dans les formules mathématiques ou est-ce l'unité servant à mesurer la valeur des variables figurant dans les fonctions, unité qui elle-même est une fonction inconnue d'une autre unité pas nommée?

Mais il se trouve, même dans les oeuvres fondamentales de l'économie mathématique, non seulement des grosses erreurs de caractère économique, mais aussi des fautes mathématiques irréparables. Leur caractéristique commun est que les auteurs ne tâchent même pas à examiner si le système d'équations accepté comme point de départ possède une solution interprétable au point de vue économique. Nous en trouvons un exemple très caractéristique dans le système d'équations connu sous le nom d'équations de WALRAS—CASSEL. Les inconnus de ces équations sont les quantités de moyens de production et de produits ainsi que leurs prix. Ce système d'équations ne peut donc avoir du sens économique que s'il a un système de solutions positives uniquement déterminé. Par contre, on a réussi à démontrer[1] que la supposition d'un tel système de solutions contient une contradiction mathématique, c'est-à-dire: le système d'équations de WALRAS—CASSEL est bien un résultat apparemment important de l'économie mathématique, mais en réalité il n'a pas de solution à laquelle on pourrait attribuer un sens économique.

Les fautes graves revenant fréquemment et systématiquement ont discrédité la science d'économie mathématique à tel point qu'aujourd'hui on peut recevoir à priori avec suspicion toute oeuvre qui traite les problèmes de l'économie politique par des méthodes mathématiques. Il serait cependant faux de proscrire tout à fait des sciences économiques les réflexions mathématiques exactes, puisque ce n'est point la méthode mathématique qui est responsable des erreurs de principe commises jusqu'ici, mais son application tout à fait erronée, ou plus précisément: l'abus commis à l'aide du formalisme mathématique. Mais l'application des résultats de la mathématique — avec la critique nécéssaire — aux phénomènes sociaux qui s'y prêtent peut conduire à des résultats précieux, parce qu'elle pourrait éclaircir plus profondément les

[1] Au sujet de la discussion mathématique concernant le système d'équations de WALRAS—CASSEL, v. Ergebnisse eines mathematischen Kolloquiums, Wien, **6** (1935), p. 10—20 et **7** (1936), p. 1—6.

rapports intrinséques des forces cachées derrière les mouvements de la société.

Ce serait naturellement une grave erreur de nous imaginer que la méthode mathématique nous rendra capables à décrire exactement les phénomènes économiques. Le développement économique est un phénomène social, il faut donc considérer à priori comme manqué toute tentative qui essayerait de le faire paraître comme un processus machinal, déterminé au sens mathématique. La méthode mathématique ne peut jamais décrire la réalité sociale toute entière, mais seulement le développement d'un modèle économique qui *dans certaines circonstances* peut approcher la réalité. Les tendances éclaircies par la mathématique économique ne peuvent donc reflêter une partie de la réalité que jusqu-au point où la structure sociale crée des situations économiques similaires au modèle supposé. Mais *la société peut changer les conditions* et alors les processus économiques réels peuvent s'écarter complètement de ceux qu'avaient établi les méthodes mathématiques, même si le modèle avait reflété primitivement la structure économique assez exactement. Il paraît que cette circonstance rend l'applicabilité des mathématiques à l'économie politique illusoire, mais en réalité c'est justement ce qui la rend légitime. C'est que les réflexions mathématiques appliquées au modèle économique correctement construit peuvent éclaircir de quels modèles la structure économique de la société doit s'écarter si nous en attendons un développement qui s'écarte *qualitativement* du développement se révélant *nécéssairement* sur le modèle. Le traitement mathématique de la science d'économie politique joue donc dans la cognition de la réalité sociale un rôle pareil par exemple à celui de la mécanique rationnelle dans la construction des appareils éléctriques: la mécanique, sans être capable de décrire complètement le fonctionnement des appareils éléctriques dont le fonctionnement est déterminé par les lois propres de l'électrodynamique, peut indiquer les difficultés mécaniques que le constructeur doit éviter s'il veut que son appareil fonctionne avec un certain effet.

L'idée fondamentale de nos investigations suivantes remonte à la théorie mathématique de VOLTERRA[2] concernant le développement des collectifs biologiques; en ce qui concerne leur contenu, elles essayent de refléter sur un modèle aussi simple que possible le développement du trafic des marchandises qui se produit sous les conditions de production du capitalisme le plus développé. Nous allons démontrer que ce développement conduit nécéssairement à ce que certains producteurs de marchandises deviennent monopolisateurs, dont il s'ensuit aisément que les conditions économiques reflétées par notre modèle rendent forcément la lutte de classes plus aiguë. Nous allons encore démontrer que l'on ne peut faire disparaître ces phénomènes même par des changements assez importants faits sur le modèle, mais seulement par

[2] L. VOLTERRA, Leçons sur la théorie mathématique de la lutte pour la vie, (Paris, 1931).

la modification complète de la structure du modèle. Exprimé d'une autre manière, cela veut dire qu'à *un certain degré du développement du capitalisme, l'intensification de la lutte de classes est inévitable si l'on conserve tant soit peu de l'essence de l'économie capitaliste.*

2. Le système d'équations différentielles du trafic des marchandises

Dans le modèle servant comme base de notre travail mathématique, nous ne faisons pas de différence entre producteurs et vendeurs de marchandises, c'est-à-dire nous ne tenons pas compte de l'intermédiaire qui s'interpose entre le producteur et le consommateur. Nous mesurons la valeur des marchandises, moyens de production etc. à un moment donné uniformément par le temps du travail socialement nécéssaire pour les produire, c'est-à-dire nous employons une mesure dont l'unité ne dépend pas des fluctuations momentanées du marché. Nos premières hypothèses sont les suivantes:

1°. La production est tellement développée que le producteur est pratiquement incapable d'utiliser ses surplus pour d'autre chose que l'augmentation de sa production, c'est-à-dire ses ventes avec profit augmentent sa production.

2°. La production de marchandises a saturé le marché au point que si le producteur A vend de la marchandise d'une valeur ξ, il y aura une quantité de marchandises de valeur $\eta = \eta(\xi)$ que le producteur B ne pourra pas placer s'il cherche à atteindre sa production maximum.

Dans les conditions du capitalisme le plus développé, ces conditions se réalisent si nous ne tenons compte que du trafic de marchandises non produites par des petits producteurs, mais seulement par des grandes installations mécaniques. Nous ajouterons que nous ne considérons comme différents deux producteurs que s'ils se font de la concurrence aiguë, c'est-à-dire si la condition sous 2° joue sur eux.

Supposons que $A_1, A_2, \ldots, A_n$ signifient tous les producteurs d'un marché qui répond à nos conditions et n'est pratiquement pas influencé par le trafic d'autres marchés. Supposons que la valeur des marchandises écoulées par eux au moment t exprimée comme fonction du temps t sera[3]

$$y_1 = y_1(t), \quad y_2 = y_2(t), \ldots, y_n = y_n(t).$$

Les producteurs emploient le surplus restant du profit provenant de la valeur des marchandises placées, conformément à notre hypothèse 1°, immédiate-

[3] Plus précisément, nous entendons par y_k la fonction définie de la manière suivante: Soit $Y_k(t)$ la valeur de la quantité totale des marchandises écoulées par le producteur A_k jusqu'au moment t; alors, y_k est la vitesse de la croissance de $Y_k(t)$, c'est à dire:

$$y_k = \frac{d\,Y_k}{dt}.$$

ment pour produire de nouvelles marchandises ce qui entraîne une augmentation du trafic de marchandises si le marché est capable d'absorber la nouvelle quantité. De cette manière, la valeur des ventes de marchandises du producteur A_k dans l'intervalle de temps $(t, t + dt)$ augmente de la quantité dy_k et cette augmentation peut être représentée comme résultante de deux composantes. L'une est la quantité que l'augmentation des ventes de marchandises de A_k atteindrait si la capacité d'absorption du marché était illimitée. Dans ce cas A_k pourrait vendre, après avoir placé une quantité de marchandises de valeur d'unité, des marchandises d'une valeur $1 + \alpha_k$ où $\alpha_k = \alpha_k(t)$ est une fonction positive ne dépendant que de l'augmentation de la capacité de production du producteur A_k résultant de l'hypothèse 1°. De cette manière, dans l'intervalle de temps $(t, t + dt)$ l'augmentation de valeur des ventes de marchandises de A_k serait

$$dy_k = \alpha_k \, y_k \, dt$$

et la valeur $y_k(t)$ serait déterminée par l'égalité

$$y_k(t) = e^{\int \alpha_k(t)\, dt}.$$

Par contre, nous avons tenir compte de l'autre composante qui résulte de la diminution de la puissance d'achat du marché découlant de notre hypothèse 2° de manière que les marchandises d'une valeur de $y_1, y_2, \ldots, y_n$ rendent impossible que A_k puisse déployer le maximum de sa capacité de vente de marchandises, ainsi la valeur des marchandises valant une unité vendues au moment t ne pourra pas accroître dans l'intervalle de temps $(t, t + dt)$ de la quantité $\alpha_k \, dt$, mais il faut réduire la quantité α_k de la valeur d'une fonction positive $\psi_k \, (y_1, y_2, \ldots, y_n)$ dépendant des variables $y_1, y_2, \ldots, y_n$; soit l'augmentation de valeur des ventes de marchandises de A_k est défini au lieu de $dy_k = \alpha_k \, y_k \, dt$ par l'équation

$$dy_k = (\alpha_k - \psi_k) \, y_k \, dt.$$

Mettons que $\varphi \, (y_1, y_2, \ldots, y_n)$ signifie la diminution totale qui se produit dans la croissance maximum $\alpha_1, \alpha_2, \ldots, \alpha_n$ des ventes de valeur unité des producteurs $A_1, A_2, \ldots, A_n$ dans l'intervalle de temps $(t, t + dt)$, alors ψ_k sera la part de la diminution φ tombant sur le producteur A_k, autrement dit

$$\psi_k \, (y_1, y_2, \ldots, y_n) = \beta_k \, \varphi \, (y_1, y_2, \ldots, y_n)$$

où $\beta_k = \beta_k(t)$ est une fonction ayant des valeurs comprises entre 0 et 1. L'équation précédente peut donc être écrite sous la forme

$$dy_k = (\alpha_k - \beta_k \, \varphi) \, y_k \, dt$$

également, c'est-à-dire qu'en désignant la dérivée de $y_k(t)$ par $y_k'(t)$, la valeur des ventes des producteurs $A_1, A_2, \ldots, A_n$ se laisse déterminer par le sys-

tème suivant d'équations différentielles:

$$
(1) \qquad
\begin{aligned}
\frac{y'_1}{y_1} &= \alpha_1 - \beta_1 \varphi(y_1, y_2, \ldots, y_n), \\
\frac{y'_2}{y_2} &= \alpha_2 - \beta_2 \varphi(y_1, y_2, \ldots, y_n), \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
\frac{y'_n}{y_n} &= \alpha_n - \beta_n \varphi(y_1, y_2, \ldots, y_n)
\end{aligned}
$$

pour toute valeur $y_k(t) \neq 0$.

3. L'existence d'un système de solutions positives

Afin de pouvoir démontrer l'existence d'un système uniquement déterminé de solutions continues et positives du système (1) d'équations différentielles de premier ordre, nous avons besoin de l'hypothèse suivante:

3°. Le marché est saturé par les quantités de marchandises vendues au point que, si n'importe quel producteur met en circulation de nouvelles marchandises, il en résulte une diminution *immédiate* de la puissance d'achat.

Cette hypothèse s'exprime en langage mathématique par l'exigence que $\varphi(y_1, y_2, \ldots, y_n)$ soit une fonction croissante en chacune de ses variables $y_1, y_2, \ldots, y_n$. Il s'ensuit immédiatement que, si $\varphi(\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n) > \varphi(y_1, y_2, \ldots, y_n)$, il existe un indice k pour lequel $\bar{y}_k > y_k$. Puisque φ mesure la diminution de la puissance d'achat du marché, il n'a de signification économique qu'en cas de valeurs bornées. Les valeurs économiquement interprétables de φ ont donc un maximum et un minimum, au-delà les valeurs de φ peuvent être définies arbitrairement. Choisissons une telle extension de φ aux valeurs économiquement non interprétables qu'il reste monotone en chacune de ses variables, de plus qu'il soit assez régulier dans le sens que, $\varepsilon > 0$ étant une quantité arbitraire, il existe toujours un $\delta > 0$, de manière que

$$\varphi(y_1, y_2, \ldots, y_{k-1}, y_k + \delta, y_{k+1}, \ldots, y_n) \geqq \varphi(y_1, y_2, \ldots, y_n) + \varepsilon .$$

Par cette extension du domaine des valeurs de la fonction φ, nous avons atteint que φ n'ait pas de maximum et qu'il ne devient infini que si au moins une des variables y_k devient également infinie. Après cela, nous démontrons le lemme suivant:

Les fonctions $y_1, y_2, \ldots, y_n$ restent bornées dans l'intervalle de temps $(0, \infty)$.

Admettons, par absurde, que y_k atteint à un certain moment une valeur supérieure à une quantité M que l'on peut choisir arbitrairement grande. Il y aura alors, en conséquence de la continuité des fonctions y_k, un moment t_M auquel

$$y_k(t_M) = M .$$

Mais si nous choisissons M assez grand, la valeur de $\varphi(y_1, y_2, \ldots, y_n)$ deviendrait, d'après ce qui précède, arbitrairement grande, quelles que soient les valeurs des fonctions y_i ayant un indice i différent de k, c'est-à-dire: même si $\overline{y_i} \leqq y_i(t_M)$ pour chaque indice $i \neq k$, on a

$$\varphi(\overline{y_1}, \overline{y_2}, \ldots, \overline{y_{k-1}}, y_k(t_M), \overline{y_{k+1}}, \ldots, \overline{y_n}) \geqq N \tag{2}$$

où N peut être arbitrairement grand, pourvu que M soit assez grand. En choisissant donc M assez grand, on peut aboutir à l'inégalité

$$N \geqq \operatorname{Max} \frac{\alpha_k}{\beta_k}. \tag{2*}$$

C'est que dans le cas contraire la relation $\frac{\alpha_k}{\beta_k} \to \infty$ devrait subsister, ce qui est impossible vu la signification économique de α_k et β_k, puisque α_k est définie par les valeurs et les frais de production des marchandises, α_k reste donc pour des raisons économiques au-dessous d'une certaine limite, quant à la valeur de β_k, elle doit rester au dessus d'une quantité positive β, sinon la diminution de la puissance d'achat du marché n'affecterait pas du tout la production de A_k, ce qui est également absurde pour des raisons d'ordre économique. La quantité $\operatorname{Max} \frac{\alpha_k}{\beta_k}$ est donc finie, il s'ensuit par conséquent d'après (2) et (2*) pour des valeurs $y_1, y_2, \ldots, y_n$ prises au moment t_M:

$$\varphi(y_1, y_2, \ldots, y_n) \geqq \frac{\alpha_k}{\beta_k}.$$

On en obtient, en vertu de (1):

$$y_k'(t_M) < 0,$$

la valeur $y_k(t)$ ne peut donc pas être $> M$ dans le voisinage de t_M. — L'hypothèse qu'après un moment t_M'

$$y_k(t) < -M'$$

où M est arbitrairement grand, conduit pareillement à une contradiction et notre lemme est entièrement démontré.

Après cela il n'est pas difficile de démontrer le théorème d'existence suivant :

Le système (1) d'équations différentielles possède en tout intervalle fini de temps un système de solutions composé de fonctions continues, positives et uniquement déterminées par leurs valeurs initiales.

Comme il ressort de (1):

$$\log \frac{y_k(t)}{y_k(t_0)} = \int_{t_0}^{t} (\alpha_k - \beta_k \varphi)\, dt.$$

Mais, selon notre lemme, $y_k(t)$ est borné et ainsi — comme nous l'avons vu — φ reste aussi sous une borne finie, donc

$$|a_k - \beta_k \varphi| \leqq C$$

où C est une constante absolue. Par conséquent

$$\left| \log \frac{y_k(t)}{y_k(t_0)} \right| \leqq C\,(t - t_0)\,,$$

c'est-à-dire:

$$y_k(t) \geqq y_k(t_0)\, e^{-C\,(t - t_0)}\,,$$

donc, si $y_k(t_0) > 0$, alors $y_k(t)$ est également positif. La détermination des fonctions $y_1, y_2, \ldots, y_k$ peut se faire par la méthode des approximations successives. La continuité et la donnée des valeurs initiales assure l'unicité du système des solutions.

4. La monopolisation du trafic de marchandises dans un système de production capitaliste très développée

Nous allons rechercher le développement du trafic des marchandises pendant une longue période tranquille de l'économie capitaliste très développée. La tranquillité de la période signifie que les conditions pour le placement des marchandises d'un producteur par rapport à celles de l'autre ne varient pas. Or les fonctions $\beta_1, \beta_2, \ldots, \beta_n$ désignent la part de la diminution totale de la puissance d'achat du marché tombant sur les producteurs $A_1, A_2, \ldots, A_n$, la tranquillité de la période se traduit donc en langage de notre symbolisme mathématique par l'hypothèse que les $\beta_1, \beta_2, \ldots, \beta_n$ sont des constantes, hypothèse que nous allons accepter pour le calcul suivant.

Éliminons φ des équations (1) et nous obtenons la relation

$$\beta_k \frac{y_i'}{y_i} - \beta_i \frac{y_k'}{y_k} = \alpha_i \beta_k - \alpha_k \beta_i$$

que nous pouvons écrire sous la forme suivante aussi:

$$d\left(\log \frac{\sqrt[\beta_i]{y_i}}{\sqrt[\beta_k]{y_k}} \right) = \left(\frac{\alpha_i}{\beta_i} - \frac{\alpha_k}{\beta_k} \right) dt.$$

Il en résulte que

$$y_k(t) = \frac{y_i(t_0)}{y_k(t_0)}\, [y_i(t)]^{\frac{\beta_i}{\beta_k}} \exp \int_{t_0}^{t} \left(\frac{\alpha_i}{\beta_i} - \frac{\alpha_k}{\beta_k} \right) dt.$$

Mais nous avons démontré au chapitre précédent que, pour un t quelconque

$0 < y_i \leqq M$, donc

$$y_k(t) \leqq \frac{y_i(t_0)}{y_k(t_0)} M^{\frac{\beta_i}{\beta_k}} \exp \int\limits_{t_0}^{t} \left(\frac{\alpha_i}{\beta_i} - \frac{\alpha_k}{\beta_k}\right) dt. \tag{3}$$

Il existe donc une constante absolue K, ainsi que

$$\frac{y_i(t_0)}{y_k(t_0)} M^{\frac{\beta_i}{\beta_k}} \leqq K . \tag{4}$$

Le quotient $\frac{\alpha_m}{\beta_m}$ exprime le rapport entre le plus grand accroissement possible des ventes de marchandises du producteur A_m et la diminution de la puissance d'achat du marché limitant cet accroissement, c'est-à-dire: $\frac{\alpha_m}{\beta_m}$ est une mesure de la capacité de vente du producteur A_m. L'expérience prouve que, s'il n'y a pas de changement essentiel dans les conditions de la concurrence, la capacité de vente du producteur qui a réussi de vendre plus pendant un temps prolongé continuera de rester plus grande que la capacité de celui dont les ventes sont restées entretemps toujours plus petites; au contraire, la différence entre les deux aura plutôt tendance à augmenter et non à diminuer. Ce fait d'expérience s'exprime dans le langage de notre formalisme mathématique que la différence $\frac{\alpha_i}{\beta_i} - \frac{\alpha_k}{\beta_k}$ ne disparaît pas si $\frac{\alpha_i}{\beta_i} - \frac{\alpha_k}{\beta_k}$ a différé de zéro pendant un temps prolongé. Si donc $\frac{\alpha_k}{\beta_k} > \frac{\alpha_i}{\beta_i}$, nous pouvons supposer par expérience l'existence d'un nombre $\gamma > 0$ ainsi que

$$\frac{\alpha_i}{\beta_i} - \frac{\alpha_k}{\beta_k} < -\gamma . \tag{5}$$

Il s'ensuit des inégalités (3), (4) et (5):

$$y_i(t) \leqq K e^{-\gamma(t-t_0)}$$

d'où nous parvenons au résultat

$$\lim_{t \to \infty} y_i(t) = 0 .$$

Cette relation est valable pour chaque indice i auquel on peut trouver un indice k de manière qu'on ait $\frac{\alpha_k}{\beta_k} > \frac{\alpha_i}{\beta_i}$ pendant un temps prolongé. Vu que, abstraction faite de quelques valeurs momentanées, on ne peut pas supposer en général $\frac{\alpha_i}{\beta_i} = \frac{\alpha_k}{\beta_k}$, parce que cela signifierait que les conditions de la con-

currence entre les producteurs A_i et A_k sont inaltérables, il n'y a, en général, qu'un seul indice k tel que $\frac{\alpha_k}{\beta_k} > \frac{\alpha_i}{\beta_i}$ pour tout i différent de k. Il n'y a donc qu'un seul producteur A_k qui soit dans la situation que la valeur $y_k(t)$ des marchandises vendues par lui reste constamment au-dessus d'une borne inférieure positive, la valeur des ventes de marchandises de tout autre producteur A_i tend, dans les conditions économiques décrites, à zéro. Nous pouvons donc énoncer le théorème suivant:

Le trafic de marchandises étant suffisamment développé et les conditions de concurrence économique assez constantes, un des producteurs tend nécéssairement à acquérir un monopole en s'accaparant tout le trafic de marchandises du marché.

Au sujet de l'applicabilité de notre théorème à la réalité, nous devons tenir compte du fait que les considérations mathématiques ne peuvent jamais décrire exactement la vie économique réelle, ils peuvent seulement éclaircir quels phénomènes devront se présenter dans la vie économique *en cas de la stabilisation de certaines conditions*. Ces conditions se reflètent dans les hypothèses faites sur le caractère du modèle mathématique. Une partie de nos hypothèses exprime la sursaturation en marchandises du marché. Ces hypothèses peuvent être acceptés comme première approximation dans la concurrence économique aiguë du capitalisme développé pour beaucoup de marchandises. L'autre partie de nos hypothèses cependant, à savoir que les fonctions $\beta_1, \beta_2, \ldots, \beta_n$ sont des constantes et l'hypothèse $\frac{\alpha_i}{\beta_i} - \frac{\alpha_k}{\beta_k} < -\gamma$ ne sont pas des concomitants indispensables du capitalisme développé. Car ces hypothèses veulent dire économiquement qu'il ne se produit pas de changement essentiel dans les conditions de concurrence des producteurs $A_1, A_2, \ldots, A_n$ par rapport de l'un à l'autre. Et de tels changements essentiels ne se produisent pas si des crises économiques éventuelles ne changent pas complètement la structure du marché. Le deuxième groupe de nos hypothèses concernant notre modèle établit donc en essence le postulat que le développement se produise sans crises économiques. Nos investigations mathématiques montrent donc que la monopolisation des marchés par des producteurs individuels ne peut pas être considerée seulement comme un *excès* du capitalisme, comme on prétend souvent à tort, mais justement au contraire: *le phénomène de la monopolisation est un concomitant aux périodes exemptes de crises du capitalisme développé. Le développement de la concurrence toujours plus aiguë du capitalisme ne mène donc pas du tout vers un nivellement des forces économiques, mais entraîne nécéssairement le développement inégal.*

Cette conclusion obtenue par des considérations mathématiques est justifiée par la supériorité sur les producteurs indépendants des trust de mam-

mouth prolifères et des cartels entrelacés par des fils compliqués. Voici comment une bonne part de ces monopoles s'est formée: les petits producteurs étaient incapables de suivre la concurrence même dans les périodes tranquilles, parce que la diminution de la capacité d'absorption du marché les aurait fait supporter des charges hors de proportion avec le bénéfice probable; c'est-à-dire l'accroissement des ventes de marchandises effectuées par le petit producteur A_i était affecté bien plus par la diminution de la capacité du marché que l'accroissement de celles effectuées par le grand producteur A_k. Ce procédé est exprimé dans le langage de notre symbolisme mathématique comme suit: la quantité β_i est grande par rapport à β_k. $\frac{\alpha_i}{\beta_i}$ est donc petit par rapport à $\frac{\alpha_k}{\beta_k}$, autrement dit: la différence $\frac{\alpha_i}{\beta_i} - \frac{\alpha_k}{\beta_k}$ reste négative, par conséquent la fonction $y_i(t)$, c'est-à-dire la valeur des ventes de marchandises du petit producteur A_i diminue exponentiellement. Nous pouvons donc considérer, en quelque sorte, le procédé de formation des monopoles connus dans la réalité comme exemple réel de notre théorie mathématique, c'est-à-dire que notre modèle mathématique reflète le développement du capitalisme moderne correctement *quant à la qualité.*

5. Le développement des ventes de marchandises effectuées par le producteur monopolisateur

En comparaison de la valeur des ventes de marchandises effectuées par le producteur A_k devenu monopolisateur à partir d'un moment t_0, nous pouvons négliger la valeur des ventes de marchandises qui pourraient éventuellement encore être effectuées par les $n - 1$ autres producteurs. A partir de ce moment, nous pouvons supprimer l'indice k et désigner par α, β et y les quantités α_k, β_k et y_k. Puisque φ aussi ne dépend que de la seule variable $y = y(t)$, le système (1) d'équations différentielles se réduit pour $t > t_0$ à la seule équation différentielle

$$\frac{y'}{y} = \alpha - \beta\, \varphi(y). \tag{6}$$

Le producteur qui se trouve dans une situation monopolisée n'est naturellement pas tenu de compter avec la force impérieuse de la concurrence, il a donc moyen de planifier sa production tant que possible et de produire tant qu'exigent les besoins du marché et, en proportion à ces besoins, le maintien et le dévéloppement de sa situation monopolisée. Le producteur monopolisateur peut donc régler sa production de manière que le développement des ventes de marchandises devienne uniforme, c'est-à-dire que le quotient $\frac{\alpha}{\beta}$ peut

être considéré pour $t > t_0$ comme constant. Dans ce cas, nous pouvons démontrer le théorème suivant:

La valeur des ventes de marchandises effectuées par le producteur monopolisé augmente à partir du moment t_0 d'une manière monotone et tend asymptotiquement vers une limite uniquement déterminée par les valeurs de α, β et φ données au moment t_0.

Admettons que y ait au moment t_1 un maximum. Dans ce cas, il s'ensuivrait de (6)

$$\varphi\,[y(t_1)] = \frac{\alpha}{\beta}$$

et pour des valeurs $t > t_1$ voisins de t_1 on devrait avoir

$$\varphi\,[y(t)] > \frac{\alpha}{\beta}\,,$$

c'est-à-dire que $\varphi(y)$ augmente autour de t_1. Or φ étant une fonction monotone de y, il s'ensuivrait $y(t) > y(t_1)$, par conséquent $y(t_1)$ ne peut pas être le maximum de la fonction $y(t)$. Par cette contradiction nous avons démontré que $y(t)$ n'a pas de maximum, mais elle est une fonction croissante pour $t > t_0$.

En ce qui concerne la deuxième partie de notre proposition, supposons que dans des intervalles assez petits $\varphi(y)$ varie proportionnellement à y et à sa valeur initiale. Il nous semble que cette hypothèse peut être acceptée comme première approximation de la réalité. Si nous introduisons donc pour simplification la notation $y_0 = y(t_0)$, $\varphi_0 = \varphi(y_0)$ et divisons l'intervalle (y_0, y) en un nombre n assez grand de parties, alors il est approximativement vrai que

$$\varphi(y) = \varphi_0 \left(1 + c\,\frac{y - y_0}{n}\right)^n$$

où c est un facteur de proportionnalité. Cette équation est d'autant plus exacte que n est plus grand, c'est-à-dire qu'ayant passé à la limite $n \longrightarrow \infty$, nous arrivons à la définition suivante de φ:

$$\varphi(y) = \varphi_0\, e^{c\,(y - y_0)}\,.$$

L'équation différentielle (6) peut donc être écrite dans la forme suivante:

$$y' = (\alpha - \beta\,\varphi_0\, e^{c\,(y - y_0)})\, y\,.$$

Nous avons cependant vu que y augmente de manière monotone restant, en même temps, fini, par conséquent c'est seulement sa valeur obtenue par le passage à la limite $t \longrightarrow \infty$ qui représente son maximum, il s'ensuit donc que

$$\lim_{t \to \infty} y'(t) = 0\,.$$

Nous en obtenons d'après la relation précédente:

$$\lim_{t \to \infty} e^{c(y-y_0)} = \frac{\alpha}{\beta \varphi_0},$$

c'est-à-dire:

$$\lim_{t \to \infty} y(t) = \frac{1}{c} \log \frac{\alpha}{\beta \varphi_0} + y_0 .$$

Cette valeur est en effet uniquement determinée par les quantités $\frac{\alpha}{\beta} = \text{const.}$, φ_0 et y_0 données au moment t_0.

6. L'influence du développement du capital monopoliste sur les salaires

Tant que la production et le trafic de marchandises sont assez équilibrés, ce que le producteur monopolisé peut atteindre par la rationalisation de sa production, la valeur de ses marchandises vendues sera proportionnelle à la valeur de l'installation et des matières premières. La valeur du capital fixe du producteur monopolisé est donc ay où a est un facteur de proportionnalité. Si le travail était gratuit, la partie fixe ay du capital augmenterait pendant le temps dt de la quantité ady. Pour pouvoir établir le profit du capitaliste, il en faut encore déduire les salaires dz revenant à la période dt comme quantité qui réduit le revenu brut. Au point de vue du capitaliste l'accroissement dx du capital est défini par

$$dx = ady - dz,$$

c'est-à-dire le capital est déterminé par l'équation fonctionnelle

$$x = ay - z .$$

Il en résulterait formellement

$$\frac{x'}{x} = \frac{ay' - z'}{ay - z};$$

nous pouvons cependant considérablement simplifier cette équation différentielle en tenant compte du fait que, selon notre théorème précédent, y croît monotonément en tendant vers une limite finie. Si nous supposons le moment t_0 assez grand, l'accroissement de y pourra être négligé pour $t > t_0$; cela revient à dire que nous pouvons supposer $y' = 0$ et $y(t) = y(t_0) = y_0$. De cette manière nous ne devons résoudre que l'équation différentielle

$$\frac{x'}{x} = - \frac{z'}{ay - z}.$$

En introduisant la notation $z_0 = z(t_0)$, nous obtenons donc le résultat

$$(7) \qquad z(t) = ay_0 - (ay_0 - z_0) \exp \int_{t_0}^{t} \frac{x'}{x}\, dt\,.$$

Vu que dans la production de la plupart de marchandises la valeur de l'installation et des matières premières à un moment donné (capital fixe) est plus grand que celui des salaires payés au même moment (capital variable), on a en général $ay_0 - z_0 > 0$. En même temps, le capitaliste tâche nécéssairement d'augmenter son capital, c'est-à-dire que $\frac{x'}{x} > 0$. Mais on a alors pour $t_2 > t_1$:

$$\int_{t_0}^{t_2} \frac{x'}{x}\, dt > \int_{t_0}^{t_1} \frac{x'}{x}\, dt.$$

Il résulte donc de (7), tenant compte de la valeur positive de $ay_0 - z_0$:

$$z(t_2) < z(t_1).$$

Nous sommes donc arrivés au résultat que, *si le capital monopoliste croît constamment, ceci entraîne — à moins que d'autres influences ne se fassent valoir — la diminution des salaires*.

Notre résultat exprime la propriété de l'époque du capitalisme de monopole que la valeur presque constante des ventes de marchandises du producteur monopolisateur *ne signifie pas* en même temps que le développement économique de la société devient uniforme, car le développement du capital monopoliste produit nécéssairement des périodes pendant lesquelles *la valeur des salaires payés par les entrepreneurs capitalistes diminue.*

Dans la réalité sociale, le rapport du développement du capital et des salaires n'est pas tellement simple comme le montre l'équation différentielle (7). C'est que le développement des salaires est influencé hors du développement du capital monopoliste examiné isolément par beaucoup d'autres facteurs sociaux que notre modèle mathématique ne reflète pas du tout. Toutefois nos raisonnements n'ont même pas le but d'exprimer la réalité sociale insaisissable par des moyens mathématiques, mais de *signaler une tendance de développement* qui dans la réalité apparaît sous une forme ou une autre; nous pouvons exprimer cette tendance le plus simplement comme suit: *il s'ensuit de la nature du développement du capital monopoliste que la lutte pour les salaires devient plus aiguë.*

7. L'apparition du processus de monopolisation en cas du règlement de la concurrence capitaliste

Nous avons tâché d'exclure de nos raisonnements, autant que possible, toute circonstance qui hors du développement du trafic de marchandises basé sur offre et demande et hors de la structure du marché pourrait encore influ-

encer la vie économique. Ainsi nous n'avons pas tenu compte par exemple des fluctuations des prix ou des possibilités offertes par le crédit et nous avons surtout négligé de prendre en considération l'influence de la structure politique de la société sur la vie économique. Malgré cela, notre modèle reflète justement les conditions du capitalisme développé, car la structure du modèle est basé sur des hypothèses économiques qui ne peuvent se réaliser que dans les conditions de production et du marché du capitalisme développé. C'est qu'une telle diminution de la capacité d'absorption du marché qu'exige notre modèle ne peut être supposée que dans une époque quand le développement téchnique de la production est capable à satisfaire la plus grande demande, par conséquent la production des marchandises dépasse les demandes moyennes du marché.

Supposant en plus certaines relations constantes par rapport de l'un à l'autre dans les capacités de vente des producteurs, nous avons également compris la concurrence aiguë du capitalisme développé parmi nos hypothèses, parce que l'exclusion de divergences essentielles ne peut être considérée comme hypothèse réelle que si la production est développée au point que le producteur ne peut pas changer essentiellement sans intervention extérieure ses rapports au marché et aux autres producteurs. Nos investigations se rapportent par conséquent aux conditions du capitalisme le plus développé; nos résultats obtenus jusqu'ici peuvent donc s'exprimer sous la forme suivante aussi: *le développement inégal créant des monopoles appartient aux propriétés intrinséques de la production capitaliste la plus développée; il en résulte nécéssairement l'apparition de certains phénomènes de la lutte pour les salaires.* Après cela, il paraît naturel de poser la question suivante: peut-on maintenir l'essence de la concurrence capitaliste, mais régler en même temps le trafic des marchandises par intervention sociale de manière qu'il ne puisse pas se développer de monopoles au détriment des autres producteurs?

Le maintien de l'essence de la concurrence capitaliste veut dire que les lois d'offre et de demande doivent en général rester en vigueur, il peut y être ajouté tout au plus encore une mesure qui restreindrait les ventes de marchandises effectuées par les producteurs individuels à point suffisant pour exclure la monopolisation du marché. L'hypothèse de la restriction se reflète mathématiquement par cela que la valeur $y_k(t)$ des ventes du producteur A_k ne pourrait augmenter de la quantité $\alpha_k y_k dt$ même si le marché était capable d'absorber n'importe combien de marchandises, parce que la mesure de restriction diminue cette augmentation d'une quantité $\lambda_k y_k\, dt$ où la grandeur de $\lambda_k = \lambda_k(t)$ dépend de la valeur de $\alpha_k = \alpha_k(t)$ et des valeurs $y_i(t)$ des ventes de marchandises effectuées des producteurs A_i différents du producteur A_k. Les fonctions $\lambda_1, \lambda_2, \ldots, \lambda_n$ sont donc des facteurs appelés à compenser le développement inégal des ventes de marchandises ayant les valeurs $y_1, y_2, \ldots, y_n$ effectuées par les producteurs $A_1, A_2, \ldots, A_n$. De cette manière, nous

obtenons, au lieu de (1), le système d'équations différentielles suivant:

$$(1a)\qquad \begin{aligned} \frac{y_1'}{y_1} &= \alpha_1 - \lambda_1 - \beta_1\,\varphi\,(y_1, y_2, \ldots, y_n), \\ \frac{y_2'}{y_2} &= \alpha_2 - \lambda_2 - \beta_2\,\varphi\,(y_1, y_2, \ldots, y_n), \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \frac{y_n'}{y_n} &= \alpha_n - \lambda_n - \beta_n\,\varphi\,(y_1, y_2, \ldots, y_n). \end{aligned}$$

Si nous admettons que les relations

$$(8)\qquad \frac{\alpha_1 - \lambda_1}{\beta_1} = \frac{\alpha_2 - \lambda_2}{\beta_2} = \ldots = \frac{\alpha_n - \lambda_n}{\beta_n}$$

subsistent pour toute valeur de t, alors les facteurs $\lambda_1, \lambda_2, \ldots, \lambda_n$ devraient changer de manière de compenser un à un les différences des quotients $\frac{\alpha_1}{\beta_1}, \frac{\alpha_2}{\beta_2}, \ldots, \frac{\alpha_n}{\beta_n}$ différents l'un de l'autre, ce qui est pratiquement irréalisable, parce qu'il est impossible de prendre des mesures qui règleraient le trafic des marchandises automatiquement. L'hypothèse (8) est par conséquent irréelle, en écrivant donc α_k' au lieu de $\alpha_k - \lambda_k$, le système $(1a)$ d'équations différentielles devient tout à fait analogue à (1), nous pouvons donc tirer de $(1a)$ les mêmes conclusions que de (1), c'est tout au plus le moment t_0 de monopolisation qui peut changer. Bien entendu, pour pouvoir raisonner ainsi, il faut supposer que les quotients $\frac{\alpha_k'}{\beta_k}$ sont doués des mêmes propriétés comme les quotients $\frac{\alpha_k}{\beta_k}$. En ce qui concerne cette question, il n'y a que trois possibilités: 1° Il existe une constante $\gamma > 0$ telle que $\frac{\alpha_i'}{\beta_i} - \frac{\alpha_k'}{\beta_k} < -\gamma$. 2° Les fonctions λ_k oscillent constamment, c'est à dire: si $\alpha_k'(t_1) > \alpha_i'(t_1)$ à un moment t_1, il existe toujours un moment $t_2 > t_1$ auquel $\alpha_k'(t_2) < \alpha_i'(t_2)$. 3° Les λ_k tendent asymptotiquement vers des valeurs déterminées par

$$\lim_{t\to\infty} \frac{\alpha_k'}{\beta_k} = \lim_{t\to\infty} \frac{\alpha_i'}{\beta_i}$$

pour tous les couples d'indices i et k.

Ce dernier cas équivaut à un règlement presque automatique du trafic des marchandises; de plus, l'automatisme doit fonctionner sans limites. Or l'hypothèse qu'un tel règlement automatique peur être imposé à la société est évidemment absurde. Dans la réalité, on ne peut donc accepter que les hypothèses 1° et 2°. Nous sommes donc parvenus au résultat suivant: si l'essence de la concurrence capitaliste est maintenue, il n'y a en réalité que deux cas possibles: *1° ou la monopolisation du trafic des marchandises et les conséquences*

qui en découlent sont inévitables malgré toutes mesures de compensation qui pourraient être introduites: 2° ou bien les facteurs de compensation causent une oscillation sans fin de la puissance de production, phénomène qui rend impossible le développement tranquille de la société. Il est donc impossible de transformer le capitalisme développé — tout en conservant son essence — en un système social dont le développement se déroule sans secousses et sans crises.

Le résultat serait tout autre si nous ne nous limitions pas uniquement à la restriction de la concurrence capitaliste, mais, en excluant complètement la possibilité de la concurrence capitaliste, nous examinerions le développement sur le modèle mathématique d'un caractère tout différent de *l'économie planifiée.* C'est que dans ce cas le développement des forces productrices et le trafic des marchandises se produit selon un plan établi d'avance; le problème mathématique ne consiste donc pas à établir le système d'équations de l'équilibre dynamique entre le développement maximum selon la téchnique de l'époque de la production des producteurs individuels d'une part et de la diminution de la puissance d'achat survenue *par suite* de ce développement d'autre part; le problème consisterait à rechercher l'optimum du développement des forces productrices en présence duquel la puissance d'achat du marché augmente au maximum. La condition préalable d'un changement si radical des conditions économiques est cependant une transformation essentielle de toute la structure de la société capitaliste dont la discussion dépasserait de loin les cadres de nos raisonnements de caractère mathématique.

(Reçu le 2 juin 1959.)

МАТЕМАТИЧЕСКАЯ ТЕОРИЯ МОНОПОЛЬ-КАПИТАЛИСТИЧЕСКОГО ТОВАРООБОРОТА

ГЕОРГИЙ АЛЕКСИЧ (Будапешт)

(Резюме)

Буржуазно-экономическая наука пыталась трактовать не мало проблем математическими методами. Попытки эти оказывались абсолютно ошибочными, потому что под видом математических форм скрывается тот взгляд, якобы капиталистическое хозяйство является единственной формой жизни. Они не выяснили, можно ли считать постоянной, единицу оборота стоимости, или она сама является функцией неизвестных факторов. Далее мы встретили даже такие трактовки, которые с математической точки зрения могут быть резко критикованы. Таким образом о системе уравнений Волрас—Касселя считавшейся в буржуазной математической экономии основной, было доказано, что она не имеет однозначно определенной положительной системы решения, т. е. математические решения системы уравнения Волрас—Касселя, не может иметь никакого экономического значения.

Мы прежде всего подчеркиваем, что математическая трактовка экономии не может быть принята без критики даже и в том случае, если она была бы лишена вышеупомянутых ошибок. Экономическое развитие есть социальное явление, всякие попытки старающиеся

доказать это развитие в математическом смысле детерминированным процессом, являются фундаментально ошибочными. Максимум чего можно достигнуть это конструкция такой модели, математические свойства которой правильно отражают одну часть экономической истины и таким образом бросают свет на то, какими силами вызываются известные исследуемые явления. На фоне этого можно определить и то, какие условия должны быть безусловно изменены обществом, чтобы элиминировать некоторые вредные явления. Цель данной статьи, что по методу исследований Вольтера в связи с биологическим коллективом, доказать на математической моделе отражающей некоторые свойства товарооборота развитого капитализма, что процесс монополизации есть неизбежное последствие развитого капитализма, и таким образом принципиально ошибочно любое представление о том, что при сохранении основных свойств капиталистического конкурса без монополизации, могло бы надеяться на спокойное экономическое развитие.

Пусть $A_1, A_2, \ldots, A_n$ означают производителей товаров определенного рынка функции $Y_1(t), Y_2(t), \ldots, Y_n(t)$ стоимость всего количества товаров, выпущенных ими до времени t и

$$y_1(t) = \frac{dY_1}{dt},\ y_2(t) = \frac{dY_2}{dt}, \ldots, y_n(t) = \frac{dY_n}{dt}.$$

Если рыночная потребность была бы безпредельна, то количество $y_k(t)$, в интервале времени $(t, t+dt)$ увеличилось бы количеством $\alpha_k y_k dt$, где $\alpha_k = \alpha_k(t)$ является положительной функцией зависящей лишь от производительной способности A_k производителя. Если предположить, что рынок настоль перегружен, что когда один из производителей выпустиг товар в оборот, тогда расположенное количество товара влечет за собой немедленное уменьшение максимально возможного оборота других производителей. В этом случае $y_k(t)$ увеличится не на $\alpha_k y_k dt$, а на $(\alpha_k - \psi_k) y_k$ где ψ_k является функцией переменных $y_1, y_2, \ldots, y_n$. Пусть $\varphi(y_1, y_2, \ldots, y_n)$ уменьшение причиненное единицей товарооборота покупной способности рынка, тогда $\psi_k = \beta_k \varphi$, где $\beta_k = \beta_k(t)$ является функцией со значениями между 0 и 1. Если $y'_k(t)$ означает производную от $y_k(t)$, тогда развитие товарооборота выражается следующей системой дифференциальных управлений:

$$\frac{y_1'}{y_1} = \alpha_1 - \beta_1 \varphi(y_1, y_2, \ldots, y_n),$$

$$\frac{y_2'}{y_2} = \alpha_2 - \beta_2 \varphi(y_1, y_2, \ldots, y_n),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\frac{y_n'}{y_n} = \alpha_n - \beta_n \varphi(y_1, y_2, \ldots, y_n).$$

Сохранивши предположение о перегруженном рынке, доказываем, что эта система дифференциальных уравнений имеет систему решения, состоящую из положительных функций, однозначно определенных начальными условиями.

Расмотрим вопрос о том возможно ли спокойное развитие товарооборота без кризиса, при производственных отношениях развитого капитализма? Такое развитие характеризовалось бы тем, что сравниваемые один с другим условия конкурса производителей $A_1, A_2, \ldots, A_n$ существенно не изменились бы, что мы можем выразить на нашей моделе рассматривая функции $\beta_1, \beta_2, \ldots, \beta_n$ постоянными. Кроме этого мы употребляем еще тот известный эмпирический факт, что способность увеличения товарооборота у A_k, за длительное время была больше, чем у A_i, тогда же в условиях развитого капитализма так и остается вообще, что на нашей

моделе выражается тем, что существует такое постоянное $\gamma > o$ для которого

$$\frac{\alpha_k}{\beta_k} - \frac{\alpha_i}{\beta_i} > \gamma.$$

На основании этих двух предположений мы доказываем, что лишь одна единственняо функция $y_k t)$ может остаться больше чем некоторое положительное настоянное, а другие асимптотически приближаются к нулю.

Экономическое обяснение этих математических данных следующее: при товарообороте соответствующем условиям развитого капитализма, и при более менее постоянных условиях конкурса, один из производителей обязательно приходит в монополизированное состояние ввиду того он охватывает совершение всего товарооборота рынка. Монополизация единичными производителями рынков, не может быть рассмотрена как болезненный прирocток капитализма, как это утверждается некоторыми буржуазными экономистами, а на самом деле, она является безусловным спутником развитого капитализма и в периодах без кризиса.

На фоне исследований товарооборота монополизированного производителя, мы доказываем, что если монополь-капитал монотонно увеличивается вместе с товарооборотом, тогда это, если другие влияния на него не действуют, влечет за собой уменьшение зарплаты.

Наконец рассмотрим вопрос о том, возможно ли сохранивши суть капиталистического конкурса, принять такие меры которые исключили бы монополизацию и ее последствия? Эта проблема математически выражается тем, что в систему дифференциальных уравнений нужно привести такие функции λ_k значение которых зависит от α_k и от значения функций y_1, y_2, $\ldots$, y_n, т. е. мы должны исследовать следующую систему дифференциальных уравнений:

$$\frac{y_1'}{y_1} = \alpha_1 - \lambda_1 - \beta_1 \varphi(y_1, y_2, \ldots, y_n),$$

$$\frac{y_2'}{y_2} = \alpha_2 - \lambda_2 - \beta_2 \varphi(y_1, y_2, \ldots, y_n),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\frac{y_n'}{y_n} = \alpha_n - \lambda_n - \beta_n \varphi(y_1, y_2, \ldots, y_n).$$

Из этих исследований мы получили тот результат, что в действительности имеются лишь две возможности: 1) или происходит процесс монополизации, несмотря на примененные принудительных мер, 2) или в товарообороте наступает такое постоянное колебание, которое полностью приводит производство в неопределенное состояние. Сохранивши суть капиталистического социального строя, становится невозможным постоянно спокойное развитие без кризиса товарооборота даже при применении принудительных мер.

TSCHIRNHAUS'SCHE EIFLÄCHEN UND EIKURVEN

Von

GYULA SZ.-NAGY (Szeged), Mitglied der Akademie

1. Eine Tschirnhaus'sche Kurve ist der Ort der Punkte der Ebene, deren Abstände von n festen Punkten der Ebene multipliziert mit gegebenen Konstanten eine konstante Summe ergeben.[1]

Eine *Tschirnhaus'sche Fläche* $E_n(R)$ *n-ten Grades mit dem Radius* R wird den Ort der Punkte P bedeuten, deren Abstände von n festen Punkten $F_1, F_2, \ldots, F_n$ *(Brennpunkte)* multipliziert mit gegebenen positiven Konstanten $q_1, q_2, \ldots, q_n$ *(Gewichte der Brennpunkte)*, deren Summe gleich n ist, die konstante Summe nR ergeben.

Liegen die Brennpunkte in einer Ebene ε, so ist die Schnittkurve der Ebene ε und der Tschirnhaus'schen Fläche $E_n(R)$ eine *Tschirnhaus'sche Kurve*. Auch diese Kurve läßt sich mit $E_n(R)$ bezeichnen. Liegen die Brennpunkte nicht in einer Ebene, so bezeichnet $E_n(R)$ nur die Fläche.

In einem multipolaren Koordinatensystem, dessen Pole die Brennpunkte sind, hat die Fläche oder Kurve $E_n(R)$ die Gleichung

$$(1) \qquad T(P) = G(P) - nR = \sum_{k=1}^{n} q_k r_k - nR = 0,$$

$$q_k > 0, \quad \sum_{k=1}^{n} q_k = n, \quad r_k = \overline{PF_k} \quad (k = 1, 2, \ldots, n).$$

Ein Punkt P liegt dann auf $E_n(R)$, wenn seine multipolaren Koordinaten $r_1, r_2, \ldots, r_n$ der linearen Gleichung (1) genügen. Eine nichtnegative Lösung $r_1, r_2, \ldots, r_n$ der Gleichung gibt aber nur dann einen Punkt von $E_n(R)$, wenn die n Kugeln von Halbmessern r_k, mit den Mittelpunkten F_k $(k = 1, 2, \ldots, n)$ einen Punkt gemeinsam haben. Dann liegt dieser Punkt

[1] Die Tschirnhaus'schen Kurven kommen erst in der Arbeit „Medicina mentis" von W. v. TSCHIRNHAUS, Amsterdam 1686, S. 91. vor. Die Tschirnhaus'sche Kurven sind spezielle Polyzonalkurven von A. CAYLEY. Vgl. G. LORIA, Spezielle algebraische und transzendente ebene Kurven, Bd I: Die algebraischen Kurven, Leipzig, 1910, S. 348–351. J. MOLNÁR hat für die Fadenkonstruktion Tschirnhaus'scher Eikurven mit ganzzahligen Gewichten ein einfaches Verfahren gegeben, *Középisk. Mat. Lapok.* 2 (1950), S. 117-121.

auf $E_n(R)$. Es wird angenommen, daß $E_n(R)$ eine reelle Fläche oder Kurve ist, daß sie also mindestens einen reellen Punkt besitzt.

Wir nennen die Flächen oder Kurven $E_n(R)$ *Tschirnhaus'sche Eiflächen* bzw. *Eikurven*, weil sie nach Satz II konvex sind.

$E_n(R)$ ist *isobar* bzw. *allgemein, anisobar*, je nachdem die Gleichungen $q_1 = q_2 = \ldots = q_n$ $(= 1)$ bestehen bzw. nicht bestehen. Die Fläche bzw. Kurve $E_1(R)$ ist eine Kugel bzw. ein Kreis vom Halbmesser R. Eine isobare bzw. allgemeine Kurve $E_2(R)$ ist eine Ellipse bzw. ein Cartesisches Oval.[2]

Haben die Brennpunkte eine symmetrische Lage in bezug auf eine Ebene, Gerade bzw. auf einen Punkt, so besitzt offenbar auch $E_n(R)$ diese Symmetrie. Eine Tschirnhaus'sche Eifläche, deren Brennpunkte auf einer Geraden g liegen, ist eine Rotationsfläche mit der Achse g. Bei einer zentralen Symmetrie der Brennpunktgruppe hat $E_n(R)$ ein Zentrum.

2. Die Flächen oder Kurven $E_n(R_1)$ und $E_n(R_2)$ sind *konfokal*, wenn sie dieselben Brennpunkte besitzen und einem Brennpunkt bei beiden Flächen oder Kurven dasselbe Gewicht gehört.

Der Radius R der konfokalen Flächen oder Kurven $E_n(R)$ hat einen von der Lage der Brennpunkte und von ihren Gewichten abhängigen kleinsten Wert R_0 mit der Eigenschaft: $E_n(R)$ ist eine reelle Fläche oder Kurve, wenn $R > R_0$ ist, im Falle $R < R_0$ hat sie keinen reellen Punkt. $E_n(R_0)$ ist ein isolierter Punkt oder eine isolierte Strecke, mit einem gemeinsamen Namen: der *Kern* der konfokalen reellen Flächen oder Kurven $E_n(R)$ $(R > R_0)$. Dieser Kern ist also ein *Kernpunkt* oder eine *Kernstrecke*.

Der Kern von $E_n(R)$ besteht aus den Punkten, in denen die Funktion $G(P) = \sum_{k=1}^{n} q_k \overline{PF_k}$ ihren kleinsten Wert $G^* = n\,R_0$ erreicht. Die stetige Funktion $G(P)$ hat in der (abgeschlossenen) konvexen Hülle H^* der Brennpunkte ein Minimum, das von G^* nicht abweichen kann. Es gibt nämlich zu jedem Punkt P außerhalb von H^* eine Ebene ε, durch welche P und H^* getrennt werden. Die orthogonale Projektion P_0 des Punktes P auf ε liegt zu jedem Punkt von H^*, also zu jedem Brennpunkt F_k, näher als P. Deshalb sind

$$\overline{P_0\,F_k} < \overline{PF_k} \quad (k = 1, 2, \ldots, n)$$

$$\text{und} \quad G^* \leqq G(P_0) = \sum_{k=1}^{n} q_k \overline{P_0\,F_k} < \sum_{k=1}^{n} q_k \overline{PF_k} = G(P)\,.$$

Der Kern von $E_n(R)$ liegt also in der konvexen Hülle H^* der Brennpunkte.

Ist H^* keine Strecke, so ist der Kern ein Kernpunkt. Die Annahme: $G(P_1) = G(P_2) = G^* = n\,R_0,\ \ P_1 \neq P_2$ führt nämlich zum Widerspruch

[2] Vgl. bei G. Loria, a. a. O., S. 173—183.

$G(H) < G^*$, wo H den Halbierungspunkt der Strecke $\overline{P_1 P_2}$ bezeichnet. Nach einem bekannten elementargeometrischen Satz besteht für einen beliebigen Punkt F die Ungleichung

$$2\,\overline{HF} \leqq \overline{P_1F} + \overline{P_2F} \tag{2}$$

und das Gleichheitszeichen gilt hier dann und nur dann, wenn F ein Punkt der Verlängerungen der Strecke $\overline{P_1P_2}$ ist (inbegriffen auch die Punkte P_1 und P_2). Enthalten also die Verlängerungen der Strecke $\overline{P_1P_2}$ nicht jeden Brennpunkt, so besteht mindestens eine der Ungleichungen

$$2\,\overline{HF_k} \leqq \overline{P_1\,F_k} + \overline{P_2\,F_k} \quad (k = 1, 2, \ldots, n) \tag{3}$$

mit Ungleichheitszeichen. Deshalb ist

$$2\,G(H) = 2\sum_{k=1}^{n} q_k\,\overline{H\,F_k} < \sum_{k=1}^{n} q_k\,\overline{P_1\,F_k} + \sum_{k=1}^{n} q_k\,\overline{P_2\,F_k} = G\,(P_1) + G(P_2) = 2\,G^*.$$

Liegen die Brennpunkte auf einer Geraden und haben sie dort die Reihenfolge $F_1, F_2, \ldots, F_n$, so gibt es einen Index p $(1 \leqq p \leqq n)$, so daß entweder

$$\sum_{k=1}^{p-1} q_k < \sum_{k=p}^{n} q_k \quad \text{und} \quad \sum_{k=1}^{p} q_k > \sum_{k=p+1}^{n} q_k, \quad \text{oder} \quad \sum_{k=1}^{p} q_k = \sum_{k=p+1}^{n} q_k$$

sind. Diese Relationen lassen sich wegen der Gleichung $\sum_{k=1}^{n} q_k = n$ in der Form

$$2\sum_{k=1}^{p-1} q_k < n < 2\sum_{k=1}^{p} q_k, \quad \text{bzw.} \quad 2\sum_{k=1}^{p} q_k = n$$

schreiben. Besteht die Ungleichung bzw. Gleichung für p, so ist der Kern von $E_n(R)$ der Kernpunkt F_p bzw. die Kernstrecke $\overline{F_p\,F_{p+1}}$, weil der Wert der Funktion $G(P)$ in F_p bzw. in den Punkten der Strecke $\overline{F_p\,F_{p+1}}$ minimal ist.

Besteht nämlich die Ungleichung für p und ist P $(\neq F_p)$ ein beliebiger Punkt der Geraden g, so ist

$$G(P) - G(F_p) = \sum_{k=1}^{n} q_k\,(\overline{P\,F_k} - \overline{F_p\,F_k}) = \sum_{k=1}^{n} q_k\,D_k > 0\,.$$

Liegt P auf der Halbgeraden $F_p\,F_1$ bzw. $F_p\,F_n$ (im Falle $p = 1$ bzw. $p = n$ ist diese Halbgerade die Verlängerung der Strecke $\overline{F_p\,F_n}$ bzw. $\overline{F_p\,F_1}$ über F_p), so erhält man aus der Reihenfolge der Brennpunkte offenbar die Relationen

$$D_k > -\overline{P\,F_p} \quad (k = 1, 2, \ldots, p-1), \quad D_k = \overline{P\,F_p} \quad (k = p,\, p+1, \ldots, n),$$

bzw.

$$D_k = \overline{P\,F_p} \quad (k = 1, 2, \ldots p), \quad D_k > -\overline{P\,F_p} \quad (k = p+1, \ldots, n)\,.$$

Daraus ergibt sich die Ungleichung

$$G(P) - G(F_p) = \sum_{k=1}^{p-1} q_k D_k + \sum_{k=p}^{n} q_k D_k > (-\sum_{k=1}^{p-1} q_k + \sum_{k=p}^{n} q_k) \overline{P F_p} > 0$$

bzw.

$$G(P) - G(F_p) = \sum_{k=1}^{p} q_k D_k + \sum_{k=p+1}^{n} q_k D_k > (\sum_{k=1}^{p} q_k - \sum_{k=p+1}^{n} q_k) \overline{P F_p} > 0 .$$

Der Punkt F_n ist also der Kernpunkt von $E_n(R)$.

Ist $\sum_{k=1}^{p} q_k = \sum_{k=p+1}^{n} q_k$, so erhält man ebenso für einen beliebigen Punkt P der Halbgeraden $F_p F_1$ bzw. $F_{p+1} F_n$ die Ungleichung $G(P) - G(F_p) > 0$ bzw. $G(P) - G(F_{p+1}) > 0$.

In einem beliebigen Punkt P der Strecke $\overline{F_p F_{p+1}}$ (F_p und F_{p+1} inbegriffen) sind

$$D_k = -\overline{PF_p} \quad (k = 1, 2, \ldots, n), \quad D_k = \overline{P F_p} \quad (k = p + 1, \ldots, n)$$

und

$$G(P) - G(F_p) = \left(-\sum_{k=1}^{p} q_k + \sum_{k=p+1}^{n} q_k\right) \overline{P F_p} = 0 .$$

Damit ist der Satz bewiesen:

I. *Der Kern einer Tschirnhaus'schen Eifläche oder Eikurve $E_n(R)$ liegt in der konvexen Hülle ihrer Brennpunkte. Liegen die Brennpunkte nicht auf einer Geraden, so ist der Kern ein Kernpunkt.*

Liegen die Brennpunkte auf einer Geraden in der Reihenfolge $F_1, F_2, \ldots, F_n$ und besteht für einen Index p die Ungleichung bzw. Gleichung

$$2 \sum_{k=1}^{p-1} q_k < n < 2 \sum_{k=1}^{p} q_k \quad \textit{bzw.} \quad 2 \sum_{k=1}^{p} q_k = n,$$

so ist der Kern von $E_n(R)$ ein Kernpunkt, der Brennpunkt F_p bzw. eine Kernstrecke, die Strecke $\overline{F_p F_{p+1}}$.

Ist also $E_n(R)$ keine Rotationsfläche, so hat sie einen Kernpunkt. Eine Rotationsfläche $E_n(R)$ hat dann eine Kernstrecke, wenn ihre Rotationsachse außerhalb der Brennpunkte einen Punkt enthält, von dem die linksseitigen Brennpunkte dieselbe Gewichtssumme ergeben, wie die rechtsseitigen.

3. Der folgende Satz ist für uns wesentlich.

II. *Jede Tschirnhaus'sche Fläche oder Kurve $E_n(R)$ $(R > R_0)$ ist eine Eifläche bzw. Eikurve.*

Dieser Satz wurde von mir zuerst mit Hilfe der Ungleichung (2) elementar bewiesen. Man kann nämlich zeigen, daß keiner der Punkte einer Sehne $\overline{P_1 P_2}$ von $E_n(R)$ $(R > R_0)$, von denen die Sehne in 2^N $(N = 1, 2, \ldots)$ gleiche Strecken geteilt werden, auf $E_n(R)$ liegen kann.

Die Konvexität von $E_n(R)$ folgt aus den Eigenschaften der konvexen Funktionen einfach.

Ist nämlich g eine Gerade durch einen beliebigen Punkt P_1 von $E_n(R)$, so ist der Abstand r_k eines auf g laufenden Punktes P vom Brennpunkt F_k eine (von unten) konvexe Funktion von P. Deshalb ist auch $G(P) = \sum_{k=1}^{n} q_k r_k$ $(q_k > 0;\ k = 1, 2, \ldots, n)$ auf g eine konvexe Funktion von P, die den Wert $G(P_1) = nR$ in höchtens einem Punkt P_2 außerhalb von P_1 annimmt. $E_n(R)$ hat also mit g höchstens zwei Treffpunkte. Daraus folgt die Konvexität von $E_n(R)$, weil ihre Punkte in einem beschränkten Raum liegen.[3]

4. Sind $R_1, R_2, \ldots, R_n$ beliebige nichtnegative reelle Zahlen, so läßt sich die Gleichung von $E_n(R)$ in der Form

$$\sum_{k=1}^{n} q_k (r_k - R_k) = nR - \sum_{k=1}^{n} q_k R_k = C \tag{4}$$

schreiben.

Bezeichnet K_h die Kugel vom Halbmesser R_h $(\geqq 0)$, deren Mittelpunkt der Punkt F_h ist, so bedeutet $d_h = r_h - R_h$ den Abstand eines Punktes P mit den multipolaren Koordinaten $r_1, r_2, \ldots, r_h$ von der Kugel K_h. Dieser Abstand ist negativ, Null oder positiv, je nachdem P innerhalb von K_h, auf K_h bzw. außerhalb von K_h liegt.

Es gilt also der Satz:

III. *Ist K_h eine Kugel von Halbmesser R_h $(\geqq 0)$, deren Mittelpunkt im Brennpunkt F_h einer Tschirnhaus'schen Fläche $E_n(R)$ liegt und hat F_h das Gewicht q_h, so genügen die Abstände d_h eines beliebigen Punktes der Fläche $E_n(R)$ von den Kugeln K_h $(h = 1, 2, \ldots, n)$ einer linearen Gleichung*

$$q_1 d_1 + q_2 d_2 + \ldots + q_n d_n = C\,.$$

Die Punkte, deren Abstände von n Kugeln mit verschiedenen Mittelpunkten einer linearen Gleichung mit positiven Koeffizienten genügen, liegen auf einer Tschirnhaus'schen Eifläche n-ten Grades, deren Brennpunkte die

[3] Ebenso kann man einsehen, daß eine reelle Fläche oder Kurve mit einer Gleichung von der Form

$$\sum_{k=1}^{n} \left[q_k^{(1)} r_k^{\nu_1} + q_k^{(2)} r_k^{\nu_2} + \ldots + q_k^{(h)} r_k^{\nu_h} + \ldots + q_k^{(m)} r_k^{\nu_m} \right] = \text{Konst} > 0,\ q_k^{(h)} \geqq 0,\ \nu_h \geqq 1$$

$(h = 1, 2, \ldots, m)$ konvex ist, weil r_h^{ν} im Falle $\nu \geqq 1$ auf einer Geraden g eine konvexe Funktion ist. Legt man nämlich g in die x-Achse, so hat r_h^{ν} die Form

$$r_k^{\nu} = f(x) = \left[(x - a_k)^2 + b_k^2 + c_k^2 \right]^{\frac{\nu}{2}}$$

und in den Punkten der x-Achse ist $f''(x) > 0$.

Kugelmittelpunkte sind. Die Gewichte der Brennpunkte sind zu den Koeffizienten der linearen Gleichung proportional.

Sind $R_1, R_2, \ldots, R_n$ die multipolaren Koordinaten eines festen Punktes P_0 von $E_n(R)$, haben also die Kugeln K_h $(h = 1, 2, \ldots, n)$ einen Punkt P_0 gemeinsam, so hat die Gleichung (4) die Form

$$(6) \qquad \sum_{h=1}^{n} q_h (r_h - R_h) = \sum_{h=1}^{n} q_h d_h = 0 , \qquad \sum_{h=1}^{n} q_h R_h = n R .$$

Daraus erhält man im Falle $n = 2$ die von I. NEWTON entdeckte Eigenschaft des Cartesischen Ovales, daß die Abstände seiner Punkte von zwei festen Kreisen in einem gegebenen Verhältnisse stehen.[4]

Die Kugeln K_h $(h = 1, 2, \ldots, n)$, deren Mittelpunkte Brennpunkte von $E_n(R)$ sind und deren Halbmesser R_h der zweiten Gleichung von (6) genügen, bilden *ein konfokales System Σ von $E_n(R)$.* Insbesondere bilden die Kugeln K_h vom Halbmesser R mit den Mittelpunkten F_h $(h = 1, 2, \ldots, n)$ ein konfokales System von $E_n(R)$. Hat die erste Gleichung von (6) ein nicht verschwindenes Glied, so hat sie mindestens je ein positives und negatives Glied.

Daraus folgt der Satz:

IV. *Ein jeder Punkt einer Tschirnhaus'schen Eifläche $E_n(R)$ liegt innerhalb einer Kugel eines konfokalen Systems Σ von $E_n(R)$ und außerhalb einer Kugel von Σ, oder er ist ein gemeinsamer Punkt der n Kugeln von Σ.*

Die Gesamtheit der gemeinsamen inneren und der Randpunkte der n Kugeln eines konfokalen Systems Σ von $E_n(R)$ enthält die Fläche. Der Durchschnitt der Kugelinneren von Σ enthält keinen Punkt von $E_n(R)$.

Aus diesem Satz erhält man im Falle $R_1 = R_2 = \ldots = R_n (= R)$ den Satz:

V. *Enthält eine Kugel K vom Halbmesser ϱ die Brennpunkte einer Eifläche $E_n(R)$ im Innern oder am Rande, so liegt jeder Punkt von $E_n(R)$ im Innern oder am Rande der konzentrischen Kugel K' vom Halbmesser $\varrho' = R + \varrho$. Im Falle $\varrho < R$ liegt die Fläche $E_n(R)$ außerhalb der konzentrischen Kugel K'' vom Halbmesser $\varrho'' = R - \varrho$.*

Die Fläche $E_n(R)$ hat bei genügend großem Radius die Form einer Kugel.

Die Kugeln vom Halbmesser R, deren Mittelpunkte in die Brennpunkte fallen, bilden nämlich ein konfokales System Σ von $E_n(R)$, sie liegen in K' und enthalten K''.

Der zweite Absatz des Satzes folgt daraus, daß das Verhältnis $\varrho' : \varrho''$ von der Einheit wenig abweicht, wenn der Radius R genügend groß ist.

[4] I. NEWTON, Philosophia naturalis, Principia mathematica, Buch I, Satz XIV. Vgl. bei G. LORIA, a. a. O., S. 178.

5. Der *Hauptpunkt* O einer Tschirnhaus'schen Eifläche oder Eikurve $E_n(R)$ ist der Schwerpunkt ihrer (mit Gewichten versehenen) Brennpunkte. Die *Hauptkugel* der Eifläche bzw. der *Hauptkreis* der Eikurve hat den Mittelpunkt im Hauptpunkt O und den Halbmesser R. Es gilt der Satz:

VI. *Eine Eifläche bzw. Eikurve $E_n(R)$ $(R > R\ )$ liegt innerhalb ihrer Hauptkugel bzw. ihres Hauptkreises und sie kann auf der Hauptkugel bzw. auf dem Hauptkreis nur dann einen Punkt haben, wenn die Brennpunkte auf einer Geraden liegen.*

Dieser Satz verallgemeinert einen bekannten Satz der Ellipse. Er scheint für ein Cartesisches Oval nicht bekannt zu sein.

Hat der Brennpunkt F_k bzw. der Hauptpunkt O in einem rechtwinkligen Koordinatensystem die Koordinaten a_k, b_k, c_k bzw. ξ, η, ζ, so sind

$$\sum_{k=1}^{n} q_k a_k = \xi \sum_{k=1}^{n} q_k = n\,\xi, \quad \sum_{k=1}^{n} q_k b_k = n\,\eta, \quad \sum_{k=1}^{n} q_k c_k = n\,\zeta. \tag{7}$$

Für einen beliebigen Punkt $P = (x, y, z)$ des Raumes bestehen also die Gleichungen

$$\begin{gathered} \sum_{k=1}^{n} q_k (a_k - x) = n (\xi - x), \quad \sum_{k=1}^{n} q_k (b_k - y) = n (\eta - y), \\ \sum_{k=1}^{n} q_k (c_k - z) = n (\zeta - z). \end{gathered} \tag{8}$$

Die linken bzw. rechten Seiten dieser Gleichungen sind Komponenten des Vektors $\sum_{k=1}^{n} q_k \overrightarrow{PF_k}$ bzw. $n\,\overrightarrow{PO}$. Die Gleichungen (8) lassen sich also in der Vektorgleichung

$$\sum_{k=1}^{n} q_k \overrightarrow{PF_k} = n\,\overrightarrow{PO} \tag{9}$$

zusammenfassen. Liegt der Punkt P auf $E_n(R)$, so ist

$$n\,R = \sum_{k=1}^{n} q_k \overline{PF_k} \geqq n\,\overline{PO}, \quad \text{also } R \geqq \overline{PO}. \tag{10}$$

Hier besteht das Gleichheitzeichen nur dann, wenn die Vektoren $\overrightarrow{PF_k}$ $(k = 1, 2, \ldots, n)$ dieselbe Richtung haben, wenn also die Brennpunkte auf der Halbgeraden PO liegen. Damit ist der Satz VI bewiesen.

6. In einem rechtwinkligen Koordinatensystem hat $E_k(R)$ die Gleichung

$$\begin{aligned} T(x, y, z) &\equiv \sum_{k=1}^{n} q_k r_k - n\,R \equiv \\ &\equiv \sum_{k=1}^{n} q_k \left[(x - a_k)^2 + (y - b_k)^2 + (z - c_k)^2\right]^{\frac{1}{2}} - n\,R = 0. \end{aligned} \tag{11}$$

Die Richtungscosinus der Normalen dieser Fläche in ihrem Punkt $P_0 = (x_0, y_0, z_0)$ sind zu $T_x(x_0, y_0, z_0)$, $T_y(x_0, y_0, z_0)$, $T_z(x_0, y_0, z_0)$ proportional.

Liegt die Normale bzw. der Punkt P_0 in der z-Achse bzw. im Anfang des Koordinatensystems und sind $d_k = (a_k^2 + b_k^2 + c_k^2)^{\frac{1}{2}}$ $(k = 1, 2, \ldots, n)$, so sind

$$-T_x(0,0,0) = \sum_{k=1}^{n} q_k \frac{a_k}{d_k} = 0, \quad -T_y(0,0,0) = \sum_{k=1}^{n} q_k \frac{b_k}{d_k} = 0,$$

$$-T_z(0,0,0) = \sum_{k=1}^{n} q_k \frac{c_k}{d_k}.$$

Der Punkt M_k, in dem die Einheitskugel mit dem Mittelpunkt P_0 von der Halbgeraden $P_0 F_k$ getroffen wird, hat die Koordinaten $x_k = \frac{a_k}{d_k}$, $y_k = \frac{b_k}{d_k}$, $z_k = \frac{c_k}{d_k}$. Wird der Punkt M_k mit dem Gewicht q_k von F_k belegt, so genügen die Koordinaten ξ_0, η_0, ζ_0 des Schwerpunktes S_0 der Punkte M_1, $M_2, \ldots, M_n$ den Gleichungen

$$\xi_0 \sum_{k=1}^{n} q_k = n\,\xi_0 = \sum_{k=1}^{n} q_k x_k = -T_x(0,0,0) = 0,$$

$$n\,\eta_0 = \sum_{k=1}^{n} q_k y_k = -T_y(0,0,0) = 0, \quad n\,\zeta_0 = \sum_{k=1}^{n} q_k z_k.$$

Der Punkt S_0 liegt also auf der Normalen von $E_n(R)$ im Punkt P_0. Daraus ergibt sich die folgende einfache Konstruktion der Normalen:

VII. *Ist M_k die zentrale Projektion des Brennpunktes F_k von einem Punkt P_0 der Eifläche $E_n(R)$ aus auf die Einheitskugel mit dem Mittelpunkt P_0, wird M_k mit dem Gewicht q_k versehen und ist S_0 der Schwerpunkt der Punkte M_1, $M_2, \ldots, M_n$, so schneidet die Gerade $P_0 S_0$ die Fläche $E_n(R)$ im Punkt P_0 senkrecht.*[5]

Aus diesem Satz folgt der Satz:

VIII. *Jede Normale einer Eifläche $E_n(R)$ $(R > R_0)$ trifft die konvexe Hülle ihrer Brennpunkte.*

7. Ein Punkt einer konvexen Fläche heißt *regulär* bzw. *singulär*, je nachdem durch ihn eine bzw. mehr als eine Stützebene der Fläche geht. Der Punkt P_0 ist also ein regulärer Punkt der Tschirnhaus'schen Eifläche $E_n(R)$, wenn die Fläche in P_0 eine bestimmte Normale besitzt. In einem singulären Punkt P_0 versagt also die eindeutige Konstruktion der Normalen. Dies findet nur

[5] Der Satz VII kommt nach G. Loria (a. a. O., S. 176) auch bei G. Peano, Applicazioni geometriche del calcolo infinitesimale, Torino 1887, S. 142 vor.

dann statt, wenn P_0 mit dem Schwerpunkt S_0 oder mit einem Brennpunkt F_k zusammenfällt, weil dann der Punkt M_k und damit S_1 unbestimmt ist.

Fallen P_0 und S_1 zusammen, so verschwinden die partiellen Differentialquotienten T_x, T_y und T_z in P_0. Enthält die Fläche $E_n(R)$ mit der Gleichung $G(P) = T(P) + nR = T(x, y, z)$ den Punkt P_0, so ist P_0 ein mehrfacher Punkt von $E_n(R)$. Wegen der Konvexität von $E_n(R)$ kann P_0 nur ein isolierter Punkt oder Punkt einer isolierten Strecke, also ein Punkt des Kerns sein. Deshalb sind $R = R_0$ und $G(P_0) = G^*$.

Die Differentialqotienten von $T(x, y, z)$ sind offenbar Komponenten des Vektors $\sum_{k=1}^{n}{}' q_k \overrightarrow{P_0 M_k}$. Daraus folgt der Satz:

IX. *Hat eine Tschirnhaus'sche Eifläche einen singulären Punkt P_0, so fällt P_0 entweder in einen Brennpunkt oder in den Kern der Fläche. In einem Punkt P_0 des Kerns besteht die Vektorgleichung* $\sum_{k=1}^{n}{}' q_k \overrightarrow{P_0 M_k} = 0$.

Aus dieser Vektorgleichung kann man auf die Lage des Kerns Folgerungen ziehen. Im Falle $q_1 = q_2 = \ldots = q_n = 1$ lassen sich die Sätze über die Lage eines Punktes kleinster Entfernungssumme von n gegebenen Punkten leicht ableiten.[6]

Die Sätze dieser Arbeit lassen sich ohne Schwierigkeiten auf die ähnlich definierten Tschirnhaus'schen Hypereiflächen mit positiven Gewichten verallgemeinern.

(Eingegangen am 10. März 1950.)

[6] H. STURM, Über den Punkt kleinster Entfernungssumme von gegebenen Punkten, *Journal für die reine und angewandte Mathematik* **97** (1884), S. 49—61. F. WEISSFELD, Sur un problème de minimum dans l'espace, *Tôhoku Math. Journal*, **42** (1936), S. 274—280.

ВЫПУКЛЫЕ ПОВЕРХНОСТИ И КРИВЫЕ ТШИРНГАУСА

Д. С.-НАДЬ (Сегед)

(Резюме)

Поверхность Тширнгауса степени n является геометрическим местом тех точек, у которых расстояние от n неподвижных точек $F_1, F_2, \ldots, F_n$ (фокусов) умноженные на данные положительные постоянные $q_1, q_2, \ldots, q_n$ (массы фокусов), сумма которых равняется n, дают постоянную сумму nR. Если фокусы лежат в одной плоскости, то кривая, которая получается как пересечение плоскости с поверхностью, называется кривой Тширнгауса. Если $q_1, q_2, \ldots, q_n$ (массы фокуса) положительны и если

$$\sum_{k=1}^{n}{}' q_k = n \quad \text{и} \quad C = nR,$$

то поверхность является выпуклой поверхностью Тширнгауса $E_n(R)$ степении n и радиуса R, так как такая поверхность выпукла.

В такой мультиполярной системе координат, где полюсами служат фокусы, уравнение поверхности имеет вид

$$\sum_{k=1}^{n} q_k r_k - nR = 0.$$

Если фокусы поверхности $E_n(R_1)$ и $E_n(R_2)$ и массы отдельных фокусов соответствует друг-другу, тогда две поверхности софокусны. Между радиусами конфокальных поверхностей $E_n(R)$ имеются такой радиус R_0, зависящий от места и веса фокусов, что поверхность $E_n(R)$ является действительной поверхностью, если $R > R_0$. Поверхности $E_n(R)$ $(R < R_0)$ не имеют действительных точек, поверхность $E_n(R_0)$ является изолированной точкой или изолированным отрезком. Последний случай только тогда имеет место, если фокусы располагаются на одной прямой и на этой прямой имеется отличная от фокуса точка, от которой сумма масс фокусов, лежащих налево от этой точки равна сумме масс фокусов, лежащих направо от этой точки.

Поверхность $E_n(R)$ шарообразна, если ее радиус достаточно велик. Поверхность $E_n(R)$ находится в том шаре с радиусом R, центр которого является центром тяжести фокусов. Одновременно $E_n(R)$ является также и геометрическим местом таких точек, у которых сумма от перемноженных с положительными массами расстояний от n шаров есть постоянная. Центрами шаров будет фокусы поверхности.

Нормаль к поверхности $E_n(R)$ в его точке P_0 строится таким образом, что фокусы проектируем на шар с центром P_0, и полученным точкам соответетствуют массы проектируемых фокусов, определяем центртяжести S_0 этих точек. Прямая $P_0 S_0$ является нормалью поверхности в точке P_0. Любая нормаль поверхности имеет общюю точку с выпуклой оболочкой фокусов. Если поверхность $E_n(R)$ пересекает один фокус, то там нет определенной нормали.

A REMARK ON THE CURVATURE AND TORTUOSITY OF SPACE-CURVES[1]

By

E. EGERVÁRY (Budapest), member of the Academy

Making a survey of the various methods of dealing with the curvature and tortuosity of a space-curve, one feels bound to ascertain that even in elementary treatises the formulae of the curvature and tortuosity are derived in a rather intricate manner, very often by the sideway of Frenet's equations.

In the present note I wish to propose a short and straightforward derivation of these formulae which rests entirely upon the definitions and requires only the most elementary tools of the vector-analysis.

The parameter t in the vector-function $\boldsymbol{v} = \boldsymbol{v}\,(t)$ being interpreted as time, it is easily proved that the angular velocity-vector of $\boldsymbol{v}\,|\,\boldsymbol{v}\,|^{-1}$ is given by $(\boldsymbol{v} \times \dot{\boldsymbol{v}})/\boldsymbol{v}^2$.

Then, having defined the curvature $1/R$ of the curve $\boldsymbol{r} = \boldsymbol{r}\,(t)$ as the limit of the ratio of the angle $\Delta\alpha$ through which the tangent-vector turns to the length Δs of the arc and identifying $\boldsymbol{v}$ with $\dot{\boldsymbol{r}}$, we get at once

$$\frac{1}{R} = \lim \frac{\Delta\alpha}{\Delta s} = \frac{d\alpha}{dt} : \frac{ds}{dt} = \frac{|\dot{\boldsymbol{r}} \times \ddot{\boldsymbol{r}}|}{|\dot{\boldsymbol{r}}|^2} : |\dot{\boldsymbol{r}}|\,.$$

Similarly, the tortuosity $1/T$ being defined as the limit of the ratio of the angle $\Delta\beta$ through which the binormal (or the osculating plane) turns to the length Δs of the arc, and identifying $\boldsymbol{v}$ with $\dot{\boldsymbol{r}} \times \ddot{\boldsymbol{r}}$ we get

$$\frac{1}{T} = \lim \frac{\Delta\beta}{\Delta s} = \frac{d\beta}{dt} : \frac{ds}{dt} = \frac{|(\dot{\boldsymbol{r}} \times \ddot{\boldsymbol{r}}) \times (\dot{\boldsymbol{r}} \times \dddot{\boldsymbol{r}})|}{|\dot{\boldsymbol{r}} \times \ddot{\boldsymbol{r}}|^2} : |\dot{\boldsymbol{r}}|\,.$$

Applying the rule $(\dot{\boldsymbol{r}} \times \ddot{\boldsymbol{r}}) \times (\dot{\boldsymbol{r}} \times \dddot{\boldsymbol{r}}) = (\dot{\boldsymbol{r}}\,\ddot{\boldsymbol{r}}\,\dddot{\boldsymbol{r}})\,\dot{\boldsymbol{r}}$, the usual expression of $1/T$ immediately follows.

[1] The author believes his method to be new, and in the contrary case he hopes to be excused by the difficulties of taking cognizance of the complete related literature.

In order to be able to distinguish the sign of the tortuosity, one has only to retain the vector-character of the foregoing expression of $\frac{d}{dt}\frac{\boldsymbol{v}}{|\boldsymbol{v}|}$.

(Received 8 May 1950)

О КРИВИЗНЕ И КРУЧЕНИИ ПРОСТРАНСТВЕННЫХ КРИВЫХ

Е. ЭГЕРВАРИ (Будапешт)

(Резюме)

Автор дает единое, прямое и краткое доказательство формулы для кривизны и кручений пространственных кривых (при любых параметрах) при помощи простых средств векторной алгебры и векторного анализа, без применения формул Френе.

ON THE REMAINDER-TERM OF THE PRIME-NUMBER FORMULA, I

By

PAUL TURÁN (Budapest), corresponding member of the Academy

1. Throughout this paper p denotes prime-numbers and $\Lambda(n)$ the classical Dirichlet-symbol

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^{\alpha}, \quad \alpha = 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

Further in the usual notation

$$\pi(x) = \sum_{p \leqq x} 1,$$

$$\psi(x) = \sum_{n \leqq x} \Lambda(n);$$

we introduce $D(x)$ and $\Delta(x)$ by

$$D(x) = \pi(x) - \int_2^x \frac{dv}{\log v} = \pi(x) - \mathrm{li}(x), \tag{1,1}$$

$$\Delta(x) = \psi(x) - x. \tag{1,2}$$

The complex variable will be $s = \sigma + it$; the zeta-function of RIEMANN is defined for $\sigma > 1$ by

$$\zeta(s) = \frac{1}{1^s} + \frac{1}{2^s} + \ldots + \frac{1}{n^s} + \ldots. \tag{1,3}$$

This function[1] is regular, apart from a pole of first order at $s = 1$, in every bounded domain. As well known, it has an infinity of zeros in $0 < \sigma < 1$ which we denote by

$$\varrho = \beta + i\gamma \tag{1,4}$$

and call non-trivial zeros; besides them $\zeta(s)$ has only the zeros

$$s = -2n, \qquad n = 1, 2, \ldots.$$

[1] All the mentioned facts on $\zeta(s)$ can be found in every standard book of the subject, we quote only INGHAM's book "The distribution of prime-numbers", Cambr. Tracts in Math. and Math. Phys., No. 30, which we shall quote often for reference.

Further we know that the non-trivial zeros are symmetrical with respect to the point $s = \frac{1}{2}$; denoting by Θ the least upper bound of the β's we have

(1,5)
$$\Theta \geqq \frac{1}{2} .$$

The significance of these zeros on the distribution of primes is signalised by the formula of RIEMANN—MANGOLDT which states that if

$$\psi(x) = \frac{\psi(x+0) + \psi(x-0)}{2} \qquad x > 1 ,$$

then we have

$$\psi(x) = x - \sum_{\varrho} \frac{x^{\varrho}}{\varrho} - \frac{1}{2} \log\left(1 - \frac{1}{x^2}\right) - \frac{\zeta}{\zeta'}(0) .$$

This shows that $\Delta(x)$ is "essentially"

(1,6)
$$= - \sum_{\varrho} \frac{x^{\varrho}}{\varrho} .$$

The classical results of HADAMARD and DE LA VALLÉE—POUSSIN according to which for $x \longrightarrow \infty$

$$\frac{\psi(x)}{x} \longrightarrow 1 , \qquad \frac{\pi(x)}{\mathrm{li}(x)} \longrightarrow 1 ,$$

raise immediately the questions on the upper and lower estimation of $\Delta(x)$ and $D(x)$. A particular emphasis is given to these questions by an assertion of RIEMANN on the distribution of primes, made at the end of his famous paper, according to which $\mathrm{li}(x)$ furnishes always an *upper* estimation of $D(x)$.

2. It is well known that we are at present much more successful with the lower estimations than with the upper ones of $\Delta(x)$ and $D(x)$. Also this paper deals exclusively with their *lower* estimation. We shall often use the notation

$$f(x) = \Omega_{\pm} (g(x))$$

where $g(x)$ is positive for all sufficiently large x-values; we denote by this notation that both inequalities

$$f(x) \geqq cg(x), \qquad f(x) \leqq -cg(x)$$

(c a suitable positive constant) have an infinity of solutions which $\longrightarrow \infty$. The symbol

$$f(x) = \Omega (g(x))$$

denotes only that the inequality

$$| f(x) | \geqq cg(x)$$

has an infinity of solutions which $\longrightarrow \infty$ for a suitable positive constant c.

3. It turned out long ago that the lower estimation of $\Delta(x)$ is easier than that of $D(x)$. My results will refer exclusively to $\Delta(x)$; the reason why

inspite of this I shall speak much in these introductory § 's also about $D(x)$ is that, by doing so, some curious features in the known results are emphasised.

It was PHRAGMÉN[1] who first succeeded in proving that if Θ denotes as before the least upper bound of the β' s, then forevery $\delta > 0$

$$\Delta(x) = \Omega_{\pm}(x^{\Theta - \delta}) \tag{3,1}$$

and ERHARD SCHMIDT[1] proved in 1903 that

$$\Delta(x) = \Omega_{\pm}(\sqrt{x}) \tag{3,2}$$

which is better than (3,1) in case $\Theta = \frac{1}{2}$. The corresponding inequality for $D(x)$ is much deeper. As ERHARD SCHMIDT remarked, if RIEMANN's assertion mentioned at the end of **1** is true, then the famous other conjecture of RIEMANN is also true, according to which all the nontrivial zeros of $\zeta(s)$ are on the line $\sigma = \frac{1}{2}$. This fact gave a new turn to the theory but it has been proved in 1914 by LITTLEWOOD[1] that

$$D(x) = \Omega_{\pm}\left(\frac{\sqrt{x}}{\log x} \log\log\log x\right). \tag{3,3}$$

This shows, of course, that the statement of RIEMANN mentioned at the end of **1** is false.

4. These results have some curious features. They do not furnish any *explicit* x-values for which (3,1), (3,2) or (3,3) is true and, in particular, for which RIEMANN's above mentioned statement is false. The reason for this curious phenomenon is principial. The proofs of the assertions (3,1), (3,2) and (3,3) use a theorem due in its ultimate form to LANDAU.[1] This states that a function defined by a series

$$\sum_{n=1}^{\infty} c_n e^{-\lambda_n s}, \qquad \lambda_1 < \lambda_2 < \ldots \to \infty$$

or an integral

$$\int_1^{\infty} c(x)\, x^{-s}\, dx$$

which is regular on the real axis for $s > \sigma_0$ but has a singular point in the half-plane $\sigma > \sigma_0$, has the property that the sequence c_ν resp. the function $c(x)$ changes its sign infinitely often. The quoted theorem states this without giving any *explicit* information about the location of these sign-changes or about the density of these. The question of density was treated by PÓLYA[2] who showed that

$$\overline{\lim_{x \to \infty}} \frac{W(x)}{\log x} > 0 \tag{4,1}$$

[2] G. PÓLYA, Über das Vorzeichen des Restgliedes im Primzahlsatz, *Gött. Nachrichten* 1930, pp. 19–27.

where $W(x)$ denotes the number of changes of sign in the sequence $\psi(n) - n$ $(n \leqq x)$. His ingenious proof is based on a general theorem, in principle similar to that of LANDAU, requiring more about the function and obtaining at that rate also conclusions of type (4,1) on the density of sign-changes. This involves that his results give no *explicit* location for these sign-changes either. In case (3,3) the matter was worsened also by the use of PHRAGMÉN-LINDELÖF's principle which in itself causes similar difficulty.

5. Therefore LANDAU's theorem and the PHRAGMÉN-LINDELÖF principle caused this indeterminedness. The latter has been completely removed by SKEWES[3] and INGHAM.[4] This was needed only in that part of LITTLEWOOD's argument which supposed $\Theta = \frac{1}{2}$, i. e., the truth of the hypothesis of RIEMANN; hence in this case they could obtain *explicit* locations for the sign-changes, in particular for the smallest value of x for which $D(x) > 0$. INGHAM even proved in this case that there is a numerical A such that for all $x > 1$ there are integers n and n' with

$$x \leqq n, n' \leqq Ax \tag{5,1}$$

and

$$D(n) > 0, \quad D(n') < 0. \tag{5,2}$$

It follows easily — denoting by $V(y)$ the number of sign-changes in the sequence $D(n)$ $(2 \leqq n \leqq y)$ — that

$$\varliminf_{y \to \infty} \frac{V(y)}{\log y} > 0 \tag{5,3}$$

which is stronger than the corresponding inequality (4,1) of PÓLYA dealing with $\Delta(x)$. INGHAM could prove (5,1) and (5,3) even in the case when $\Theta > \frac{1}{2}$, supposing that there is a zero on the line $\sigma = \Theta$, but of course his lower estimation depends then on this hypothetical zero: if the line $\sigma = \Theta$ does not contain a ϱ, his method fails at the present. The inequality which would correspond to (5,1) for $\Delta(x)$ instead of $D(x)$ is nowhere explicitly stated; the other one corresponding to (5,3) is proved by PÓLYA[2] supposing again $\Theta = \frac{1}{2}$.

6. Thus we have a rather intricate situation. All difficulties arise obviously from the fact that the lower estimation is made dependent upon the

[3] S. SKEWES, On the difference $\pi(x) - \mathrm{li}(x)$, *Journal of Lond. Math. Soc.*, 8 (1933), pp. 277—283. I learn from a letter of Prof. LITTLEWOOD that Mr. SKEWES will publish a paper containing a complete solution of the problem of the least sign-change, giving without any supposition an *explicit* numerical upper bound for it. (Letter dated from 29. Aug. 1949.)

[4] A. E. INGHAM, A note on the distribution of primes, *Acta Arith.*, 1 (1936), pp. 201—211.

configuration of *all* non-trivial zeros, about which our knowledge is rather poor. As LITTLEWOOD[5] wrote „Those familiar with the theory of the RIEMANN zeta-function in connection with the distribution of primes may remember that the interference difficulty arises with the function

$$f(x) = \sum_{\varrho} \frac{x^{\varrho}}{\varrho} = \sum \frac{x^{\beta + i\gamma}}{\beta + i\gamma}$$

(where the ϱ's are the complex zeros of $\zeta(s)$). There exist proofs that if Θ is the upper bound of the β's (so that $\Theta = \frac{1}{2}$ if RIEMANN hypothesis is true) then $f(x)$ is of order at least $x^{\Theta - \varepsilon}$ in x. But these proofs are curiously indirect: if $\left(\Theta > \frac{1}{2}\right.$ and$\left.\right)$ we are given a particular $\varrho = \varrho_0$ for which $\beta = \beta_0 > \frac{1}{2}$, they provide no explicit X depending only upon β_0, γ_0 and ε such that $|f(x)| > X^{\beta_0 - \varepsilon}$ for some x in $(0, X)$." In other words, he asks for the Ω-estimation of $D(x)$ and $\Delta(x)$ to replace LANDAU's theorem by another estimation *depending on one single ϱ only, instead of one depending on the „smallest" half-plane containing all the zeros ϱ*. And he adds „There is no known way of showing (for any explicit X) that the single term $\frac{x^{\beta_0 + i\gamma_0}}{\beta_0 + i\gamma_0}$ of f is not interfered with by the other terms of the series over the range $(0, X)$". In what follows I shall show in the case of $\Delta(x)$ such a way which is workable also for many similar problems and means a new approach to the problem. This way consists in the application of a method which I applied several times for various aims in the analysis and analytical theory of numbers.[6] If $c_0, c_1, \ldots$ denote either numerical constants whose values can be given *explicitly*, or *explicitly given* functions of ϱ_0 (the latter case will be always stated explicitly), then I shall prove the following

THEOREM. If $\varrho_0 = \beta_0 + i\gamma_0$, $\beta_0 > \frac{1}{2}$ is an arbitrary zero of $\zeta(s)$, then for[7]

$$T > \max\,(c_0, c_1(\varrho_0))$$

we have

$$\max_{1 \leq x \leq T} |\Delta(x)| > \frac{T^{\beta_0}}{|\varrho_0|^{10 \frac{\log T}{\log\log T}}} e^{-c_{17} \frac{\log T \log\log\log T}{\log\log T}}.$$

[5] J. E. LITTLEWOOD, Mathematical notes (12). An inequality for a sum of cosines. *Journ. of Lond. Math. Soc.*, **12** (1937), pp. 217—222.

[6] For a discussion of the applications see my lecture at the Meeting of the Czechoslovakian and Polish Mathematical Associations in Prague delivered on 3. Sept. 1949 entitled "On a new method in the analysis with applications".

[7] For c_0 we have only to choose max $(c_5, c_{12}, c_{13}, c_{14})$; for $c_1(\varrho_0)$ we have only the requirements (15.2) and (16.5).

7. Hence taking[8] e. g. $\varrho_0 = \frac{1}{2} + i\ 14{,}13 \ldots$ we obtain that for $T > c_2$ we have *without any supposition*

$$\max_{1 \leq x \leq T} |\Delta(x)| > \sqrt{T}\ e^{-c_3 \frac{\log T \log\log\log T}{\log\log T}}. \tag{7,1}$$

8. The main tool in the proof of the theorem will be the following

LEMMA[9]. If $n \leq N$, $m \geq 28\ N$ and the complex numbers $z_1, z_2, \ldots, z_n$ are subjected only to the condition

$$\max_{\nu = 1, 2, \ldots, n} |z_\nu| \geq 1, \tag{8,1}$$

then we have

$$\max_{m \leq \nu \leq m+N} |z_1^\nu + z_2^\nu + \ldots + z_n^\nu| \geq \frac{1}{N^2} \left(\frac{N}{e^{27}\, m}\right)^N. \tag{8,2}$$

9. INGHAM's theorem (5,1) — (5,2) suggests among others the question: „how often" can $\Delta(x)$ be large positively and negatively without any hypothesis? My method can not answer to this question at this moment; it would need instead of the above formulated lemma another one giving one-sided estimations for $R(z_1^\nu + z_2^\nu + \ldots + z_n^\nu)$. My original lemma admits an estimation, analogous to (8,2), also for expressions $(b_1\, z_1^\nu + \ldots + b_n\, z_n^\nu)$, but I have shown by counter-examples[10] that there is no one-sided theorem in such a generality even if we suppose all the coefficients b_j to be positive. But information can easily be derived for $|\Delta(x)|$. As a matter of fact, since for $T > 1$

$$\max_{1 \leq x \leq T^{\frac{3}{10}}} |\Delta(x)| < 10\ T^{\frac{3}{10}},$$

we obtain for $T > c_4$

$$\max_{T^{\frac{3}{10}} \leq x \leq T} |\Delta(x)| > T^{\frac{4}{10}}. \tag{9,1}$$

No doubt, my argument can be modified to tighten the interval $(T^{\frac{3}{10}}, T)$, even perhaps to an interval $(T - h(T), T)$, where $h(T) = o(T)$. Such a result could be easily deduced from a mean-value theorem for $|\Delta(x)|$ or $\Delta(x)^2$; but such one exists only in the case when RIEMANN's hypothesis is true.[11]

[8] See e. g. E. C. TITCHMARSH, The Zeta-Function of Riemann, Cambr. Tracts in Math. and Math. Phys., No. 26, p. 45.

[9] P. TURÁN, On Riemann's hypothesis, *Bull. de l'Acad. des Sciences de l'URSS*, **11** (1947), pp. 197—262.

[10] P. TURÁN, On a theorem of Littlewood, *Journ. of the Lond. Math. Soc.*, **21** (1946), pp. 268—275.

[11] H. CRAMÉR, Ein Mittelwertsatz in der Primzahltheorie, *Math. Zeitschr.*, **12** (1946), pp. 147—153.

10. As well known[1]

$$\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = -\frac{\zeta'}{\zeta}(s) ;$$

hence from (1,3)

$$(10,1) \qquad \sum_{n=1}^{\infty} \frac{\Lambda(n)-1}{n^s} = -\left(\frac{\zeta'}{\zeta}(s) + \zeta(s)\right) = F(s).$$

Since $\zeta(s)$ has a pole of first order at $s = 1$, it follows easily that $F(s)$ is regular for $s = 1$; thus $F(s)$ is singular only at $s = \varrho$ and $s = -2n$ $(n = 1, 2, \ldots)$, where all singularities are poles of first order with residua 1. Let T be so large that

$$\frac{20}{\log\log T} < \frac{3}{2}, \qquad T > 100, \qquad 44 \log^2 T < e^{41 \frac{\log T \log\log\log T}{\log\log T}},$$

$$(10,2) \qquad 9 < \log^{\frac{1}{30}} T < \log^{\frac{1}{10}} T (\log\log T)^3 < \frac{\log^{0,9} T}{\log\log T} < 10 \frac{\log T}{(\log\log T)^2},$$

$$20 \log^{\frac{1}{10}} T < T, \quad \log^2 (1 + \log^{\frac{1}{10}} T) < \log^{\frac{1}{10}} T;$$

all of these are satisfied if

$$T > c_5.$$

Let K_0 and N_0 be defined by

$$(10,3) \qquad K_0 = 10 \frac{\log T}{\log\log T},$$

$$(10,4) \qquad N_0 = \log^{\frac{1}{10}} T (\log\log T)^2.$$

From (10,2) we have

$$(10,5) \qquad K_0 > N_0 > 9.$$

Then there is an $L > 2$ for which

$$(10,6) \qquad (L^{K_0} <) L^{K_0+N_0} \leqq T < L^{K_0+N_0+1} (< L^{3K_0});$$

for this L we have

$$(10,7) \qquad (9 <) \log^{\frac{1}{30}} T \leqq L \leqq \log^{\frac{1}{10}} T.$$

After fixing these, we restrict the quantities U and the integer k at this moment only by

$$(10,8) \qquad L \leqq U < L + 1$$

and

$$(10,9) \qquad (9 <) K_0 < k + 1 < K_0 + N_0 (< 2 K_0).$$

11. First we consider the integral

$$(11,1) \qquad J(T) = \frac{1}{2\pi i} \int_{1+\eta-iU}^{1+\eta+iU} \frac{\xi^s}{s^{k+1}} F(s)\, ds$$

where

$$(11,2) \qquad \xi = L^{k+1}, \qquad \eta = \frac{1}{\log (L^{k+1})}.$$

Using the representation (10,1) which converges absolutely, we have

$$(11,3) \qquad J(T) = \sum_{n=1}^{\infty} (\Lambda(n) - 1) \frac{1}{2\pi i} \int_{1+\eta-iU}^{1+\eta+iU} \frac{\left(\frac{\xi}{n}\right)^s}{s^{k+1}} ds .$$

If we write the integral in (11,3) in the form

$$\frac{1}{2\pi i} \int_{1+\eta-i\infty}^{1+\eta+i\infty} - \frac{1}{2\pi i} \int_{1+\eta-i\infty}^{1+\eta-iU} - \frac{1}{2\pi i} \int_{1+\eta+iU}^{1+\eta+i\infty},$$

then for the first we have the classical value

$$\frac{1}{k!} \log^k \frac{\xi}{n} \qquad \text{if } n \leqq \xi$$

and

$$0 \quad \text{if} \quad n > \xi;$$

for the second and third we have trivially the upper bound

$$\frac{1}{2\pi} \left(\frac{\xi}{n}\right)^{1+\eta} \int_U^{\infty} \frac{dt}{t^{k+1}} = \frac{e}{2\pi} \frac{\xi}{n^{1+\eta}} \frac{1}{k\, U^k} \leqq \frac{e}{2\pi} \frac{L}{k\, n^{1+\eta}}.$$

Collecting these we obtain from (11,3)

$$\left| J(T) - \sum_{n \leqq \xi} (\Lambda(n) - 1) \frac{\log^k \frac{\xi}{n}}{k!} \right| < \frac{e}{\pi} \frac{L}{k} \sum_{n=1}^{\infty} \frac{\log n}{n^{1+\eta}} < \frac{L}{k} \left(1 + \frac{1}{\eta^2}\right) =$$

$$= \frac{L}{k} \{1 + (k+1)^2 \log^2 L\} < \frac{L}{k} \{1 + 2\, k^2 \log^2 L\} < 3\, kL \log^2 L$$

or using (10,7), (10,9), (10,3) and (10,2)

$$< 3.20 \frac{\log T}{\log \log T} \log^{\frac{1}{10}} T \frac{1}{10^2} (\log \log T)^2 < \log^2 T.$$

Or

$$(11,4) \qquad | J(T) | \leqq \frac{1}{k!} \sum_{n \leqq \xi} (\Lambda(n) - 1) \log^k \frac{\xi}{n} + \log^2 T.$$

Since we have

$$\Delta(1) - 1 = \Delta(1)$$

and for $n \geqq 2$

$$\Delta(n) - 1 = \Big\{\sum_{\nu \leqq n} \Lambda(\nu) - n\Big\} - \Big\{\sum_{\nu \leqq n-1} \Lambda(\nu) - (n-1)\Big\} = \Delta(n) - \Delta(n-1),$$

we may write the sum in (11,4) in the form

$$\Delta(1)\log^k \frac{\xi}{1} + (\Delta(2) - \Delta(1))\log^k \frac{\xi}{2} + \ldots + \{\Delta([\xi]) - \Delta([\xi]-1)\}\log^k \frac{\xi}{[\xi]} =$$

$$= \Delta(1)\left(\log^k \frac{\xi}{1} - \log^k \frac{\xi}{2}\right) + \Delta(2)\left(\log^k \frac{\xi}{2} - \log^k \frac{\xi}{3}\right) + \ldots$$

$$+ \Delta([\xi]-1)\left(\log^k \frac{\xi}{[\xi]-1} - \log^k \frac{\xi}{[\xi]}\right) + \Delta([\xi])\left(\log^k \frac{\xi}{[\xi]} - \log^k \frac{\xi}{[\xi]}\right) =$$

$$= -\int_1^{\xi} \Delta(x) \frac{d}{dx}\left(\log^k \frac{\xi}{x}\right) dx.$$

Hence (11,4) takes the form

$$|J(T)| \leqq \log^2 T + \frac{1}{k!}\left|\int_1^{\xi} \Delta(x)\frac{d}{dx}\left(\log^k \frac{\xi}{x}\right) dx\right|$$

or a fortiori

(11,5) $$|J(T)| \leqq \frac{1}{k!}\int_1^{\xi} |\Delta(x)| \left|\frac{d}{dx}\log^k \frac{\xi}{x}\right| dx + \log^2 T.$$

12. To find a lower estimation for $|J(T)|$, we shall use first in a usua way CAUCHY's theorem. It is well known[1] that for $m = 2, 3, \ldots$ there is a T_m with

(12,1) $$m < T_m < m+1$$

and

(12,2) $$\left|\frac{\zeta'}{\zeta}(s)\right| < c_6 \log^2 T_m, \qquad -1 \leqq \sigma \leqq 2.$$

Since the integer L is > 2, according to (10,7) we determine U in the integral (11,1) by

(12,3) $$U = T_L$$

and we apply CAUCHY's theorem to the function

$$\frac{\xi^s}{s^{k+1}} F(s)$$

along the parallelogram l formed by the segments l_1, l_2, l_3, l_4 where these are defined by

$$\sigma = 1 + \eta, \qquad - T_L < t < T_L,$$

$$t = T_L, \qquad - 2\left[\frac{k}{2}\right] + 1 \leqq \sigma \leqq 1 + \eta,$$

$$t = - T_L, \qquad - 2\left[\frac{k}{2}\right] + 1 \leqq \sigma \leqq 1 + \eta,$$

$$\sigma = - 2\left[\frac{k}{2}\right] + 1, \qquad - T_L \leqq t \leqq T_L$$

respectively. According to what was said in **10** after the formula (10,1) we have

$$\frac{1}{2\pi i}\int_{(l)} \frac{\xi^s}{s^{k+1}} F(s)\, ds = \sum_{|\gamma| < T_L} \frac{\xi^\varrho}{\varrho^{k+1}} + \frac{1}{k!} \left(\xi^s F(s)\right)^{(k)}_{s=0} + \sum_{n=1}^{\left[\frac{k}{2}\right]-1} \frac{\xi^{-2n}}{(-2n)^{k+1}}$$

or

$$|J(T)| > \left|\sum_{|\gamma| < T_L} \frac{\xi^\varrho}{\varrho^{k+1}}\right| - \frac{1}{k!}\left|\left(\xi^s F(s)\right)^{(k)}_{s=0}\right| - \left|\sum_{n=1}^{\left[\frac{k}{2}\right]-1} \frac{\xi^{-2n}}{(-2n)^{k+1}}\right| - \tag{12,4}$$

$$- \frac{1}{2\pi}\left|\int_{(l_2)}\right| - \frac{1}{2\pi}\left|\int_{(l_3)}\right| - \frac{1}{2\pi}\left|\int_{(l_4)}\right| = |J_1| - |J_2| - |J_3| - |J_4| - |J_5| - |J_6|.$$

13. We need upper estimations for $|J_2|$, $|J_3|$, ..., $|J_6|$. All these are routine work. For J_3 we have simply

$$|J_3| < \xi^{-2} = L^{-2(k+1)} \leqq 1. \tag{13,1}$$

To estimate J_2 we first remark that

$$J_2 = \frac{1}{k!}\sum_{\nu=0}^{k} \binom{k}{\nu} \log^\nu \xi \, F^{(k-\nu)}(0),$$

$$|J_2| \leqq \sum_{\nu=0}^{k} \frac{1}{\nu!\,(k-\nu)!} \log^\nu \xi \,|F^{k-\nu}(0)|. \tag{13,2}$$

Using the known fact[8] that $\zeta(s)$ does not vanish in the circle $|s| \leqq \frac{3}{2}$, it follows from **10** that $F(s)$ is regular in the circle $|s| \leqq \frac{3}{2}$ and here

$$|F(s)| \leqq c_7. \tag{13,2}$$

Hence CAUCHY's coefficient-estimation gives

$$|F^{(k-\nu)}(0)| \leqq c_7\,(k-\nu)!\left(\frac{2}{3}\right)^{k-\nu}.$$

Putting this into (13,2) we obtain

$$|J_2| \leqq c_7 \left(\frac{2}{3}\right)^k \sum_{\nu=0}^{k} \frac{1}{\nu!} \left(\frac{3}{2} \log \xi\right)^\nu;$$

using (10,9) and (10,6) we have

$$\log \xi = \log (L^{k+1}) \leqq \log (L^{K_0+N_0}) \leqq \log T.$$

Hence

$$(13,4) \qquad |J_2| \leqq c_7 \left(\frac{2}{3}\right)^k \sum_{\nu=0}^{k} \frac{1}{\nu!} \left(\frac{3}{2} \log T\right)^\nu.$$

But the sequence

$$\frac{1}{\nu!} \left(\frac{3}{2} \log T\right)^\nu \qquad \nu = 0, 1, \ldots$$

increases for $\nu \leqq \frac{3}{2} \log T$; this means that — owing to (10,9) — the terms of the sum in (13,4) are certainly increasing if

$$\frac{20 \log T}{\log \log T} = 2 K_0 \leqq \frac{3}{2} \log T$$

which is true according to (10,2). Hence from (10,9)

$$|J_2| \leqq c_7 \left(\frac{2}{3}\right)^k (k+1) \frac{1}{k!} \left(\frac{3}{2} \log T\right)^k < c_7 \frac{2 K_0}{k!} \log^k T =$$

$$= 20\, c_7 \frac{\log T}{\log \log T} \frac{\log^k T}{k!} < 20\, c_7 \log T \left(\frac{e \log T}{k}\right)^k < 20\, c_7 \log T \left(\frac{6 \log T}{K_0}\right)^{2K_0} =$$

$$(13,5) \qquad = 20\, c_7 \log T \left(\frac{3}{5} \log \log T\right)^{20 \frac{\log T}{\log \log T}} < c_7\, e^{41 \frac{\log T \log \log \log T}{\log \log T}}$$

using also (10,2). Hence from (12,4), (13,1) and (13,5) we have

$$(13,6) \qquad |J(T)| > |J_1| - 1 - c_7\, e^{41 \frac{\log T \log \log \log T}{\log \log T}} - |J_4| - |J_5| - |J_6|.$$

14. For the estimation of $|J_4|$, $|J_5|$ and $|J_6|$ we need the well-known inequalities[8]

$$(14,1) \qquad \left|\frac{\zeta'}{\zeta}(s)\right| \leqq c_8 \log |s| \leqq c_8 |s|$$

valid for the half-plane $\sigma \leqq -1$ removing from it the circles

$$|s - n| < \frac{1}{2}, \quad (n = -2, -4, \ldots),$$

and

$$(14,2) \qquad |\zeta(s)| \leqq c_9 \{1 + (|t| + 2)^{1-\sigma}\} \log (|t| + 2),$$

$$(14,3) \qquad |\zeta(s)| \leqq c_9 \{(|t| + 2) + (|t| + 2)^{1-\sigma} \log (|t| + 2)\}$$

valid uniformly for $|s-1| \geqq 1$. Hence from (14,1) and (14,3)

$$|J_6| \leqq \xi^{1-2\left[\frac{k}{2}\right]} \left(\int_{-(L+1)}^{L+1} \frac{c_8 |s|}{|s|^{k+1}} dt + c_9 \int_{-(L+1)}^{L+1} \frac{(|t|+2) + (|t|+2)^k \log(|t|+2)}{\left(\frac{|t|+2}{\sqrt{2}}\right)^{k+1}} dt \right) <$$

$$< \xi^{3-k}\, 2^{\frac{k+1}{2}} \left\{ 2 c_8 \int_0^\infty \frac{dt}{1+t^2} + 2 c_9 \int_0^\infty \frac{dt}{(2+t)^2} + 2 c_9 \int_0^{L+1} \frac{\log(t+2)}{t+2} dt \right\} <$$

$$< 4 \left(\frac{2}{\xi}\right)^{\frac{k+1}{2}} \{ c_8 + c_9 + c_9 (L+3) \} < 4 L^{-\frac{(k+1)^2}{4}} \{ c_8 + 4 c_9 + c_9 \log^{\frac{1}{10}} T \} .$$

But

$$\frac{(k+1)^2}{4} \geqq \frac{K_0^2}{4} \geqq 2 K_0 > K_0 + N_0 + 1 ,$$

i. e.,

$$(14,4) \qquad |J_6| < \frac{4}{T} (c_8 + 4 c_9 + c_9 \log^{\frac{1}{10}} T) < (c_8 + c_9) \frac{20 \log^{\frac{1}{10}} T}{T} < (c_8 + c_9)$$

owing to (10,6) and (10,2). To estimate $|J_4|$ and $|J_5|$ it is obviously sufficient to consider $|J_4|$ only. For that part of l_2 on which $\sigma \leqq -1$ we have again

$$(14,5) \qquad |F(s)| \leqq c_8 |s| + c_9 \{ |t| + 2 + (|t|+2)^{1-\sigma} \log(|t|+2) \} ;$$

for the remaining part of l_2 we have, using (12,2) and (14,3),

$$|F(s)| \leqq c_6 \log^2 T_L + c_9 \{ |t| + 2 + (|t|+2)^{1-\sigma} \log(|t|+2) \} <$$
$$< c_6 |s| + c_9 \{ |t| + 2 + (|t|+2)^{1-\sigma} \log(|t|+2) \} \cdot$$

i. e., again the estimation (14,5) but with c_6 instead of c_8. Replacing both c_6 and c_8 by $(c_6 + c_8)$ we have on the whole segment l_2

$$(14,6) \quad |F(s)| \leqq (c_6 + c_8) |s| + c_9 \{ |t| + 2 + (|t|+2)^{1-\sigma} \log(|t|+2) \} .$$

Hence

$$|J_4| < \int_{-\infty}^{1+\eta} \xi^\sigma \left\{ \frac{c_6 + c_8}{|s|^k} + c_9 \frac{L + 2 + (L+2)^{1-\sigma} \log(L+2)}{L^{k+1}} \right\} d\sigma <$$

$$< \frac{c_6 + c_8}{L^k} (L^{k+1})^{1+\eta} + \frac{2 c_9}{L^k} (L^{k+1})^{1+\eta} + c_9 \frac{(L+2)\log(L+2)}{L^{k+1}} \int_{-\infty}^{1+\eta} \left(\frac{\xi}{L+2}\right)^\sigma d\sigma <$$

$$< e (c_6 + c_8 + 2 c_9) L + \frac{2 c_9 \log(L+2)}{L^k} \int_{-\infty}^{1+\eta} L^{k\sigma} d\sigma < 3 (c_6 + c_8 + 2 c_9) \log^{\frac{1}{10}} T +$$

$$+ \frac{4 c_9 \log L}{L^k} \cdot \frac{L^k \cdot e}{\log L} < 3 (c_6 + c_8 + 2 c_9) \log^{\frac{1}{10}} T + 12 c_9 <$$

$$(14,7) \qquad < 3\,(c_6 + c_8 + 6\,c_9)\,\log^{\frac{1}{10}} T.$$

Collecting (13,6), (14,4) and (14,7), we have owing to (10,2)

$$|J(T)| > |J_1| - 1 - c_7\, e^{41\frac{\log T \log\log\log T}{\log\log T}} - 6\,(c_6 + c_8 + 6\,c_9)\log^{\frac{1}{10}} T -$$

$$- (c_8 + c_9) > |J_1| - 7\,(c_6 + c_7 + c_8 + 6\,c_9 + 1)\,e^{41\frac{\log T \log\log\log T}{\log\log T}} =$$

$$= |J_1| - c_{10}\, e^{41\frac{\log T \log\log\log T}{\log\log T}}.$$

Putting this into (11,5) we have

$$(14,8) \qquad \begin{aligned} |J_1| &< c_{10}\, e^{41\frac{\log T \log\log\log T}{\log\log T}} + \log^2 T + \frac{1}{k!}\int_1^{\xi} |J(x)| \left|\frac{d}{dx}\log^k\frac{\xi}{x}\right| dx < \\ &< (c_{10} + 1)\, e^{41\frac{\log T \log\log\log T}{\log\log T}} + \frac{1}{k!}\int_1^{\xi} |J(x)| \left|\frac{d}{dx}\log^k\frac{\xi}{x}\right| dx . \end{aligned}$$

15. To obtain finally a *lower* estimation of $|J_1|$ by a suitable choice of k we write it first in the form

$$|J_1| = \frac{\xi^{\beta_0}}{|\varrho_0|^{k+1}}\left|\sum_{|\gamma|<T_L} \xi^{\varrho-\varrho_0}\left(\frac{\varrho_0}{\varrho}\right)^{k+1}\right| > \frac{L^{(k+1)\beta_0}}{|\varrho_0|^{k+1}}\left|\sum_{|\gamma|<T_L}\left(L^{\varrho-\varrho_0}\frac{\varrho_0}{\varrho}\right)^{k+1}\right| =$$

$$(15,1) \qquad = \left(\frac{L^{\beta_0}}{|\varrho_0|}\right)^{k+1}\left|\sum_{|\gamma|<T_L}\left(L^{\varrho-\varrho_0}\frac{\varrho_0}{\varrho}\right)^{k+1}\right|.$$

Let T be so large that

$$(15,2) \qquad \log^{\frac{1}{60}} T > |\varrho_0| .$$

Then we have a fortiori from (10,7)

$$L^{\beta_0} \geqq (\log^{\frac{1}{30}} T)^{\frac{1}{2}} > |\varrho_0| ,$$

i. e., from (10,6), (10,3) and (10,7)

$$\left(\frac{L^{\beta_0}}{|\varrho_0|}\right)^{k+1} > \left(\frac{L^{\beta_0}}{|\varrho_0|}\right)^{K_0} = \frac{L^{\beta_0(K_0+N_0+1)}}{|\varrho_0|^{K_0}}\, L^{-\beta_0(N_0+1)} >$$

$$> \frac{T^{\beta_0}}{|\varrho_0|^{10\frac{\log T}{\log\log T}}}\, \frac{1}{(\log^{\frac{1}{10}} T)^{2\log^{\frac{1}{10}} T\,(\log\log T)^2}} = \frac{T^{\beta_0}}{|\varrho_0|^{10\frac{\log T}{\log\log T}}}\, e^{-\frac{1}{5}\log^{\frac{1}{10}} T\,(\log\log T)^2} .$$

Thus from (15,1)

$$(15,3) \qquad |J_1| > \frac{T^{\beta_0}}{|\varrho_0|^{10\frac{\log T}{\log\log T}}}\, e^{-\frac{1}{5}\log^{\frac{1}{10}} T\,(\log\log T)^2}\left|\sum_{|\gamma|<T_L}\left(L^{\varrho-\varrho_0}\frac{\varrho_0}{\varrho}\right)^{k+1}\right| .$$

16. We denote the remaining sum by Z. We shall use now our lemma formulated in **8.** Since L and T_L are defined as functions of T only, the number of the terms and also the terms $\left|L^{\varrho-\varrho_0}\frac{\varrho_0}{\varrho}\right|$ themselves are independent of k; they will play the role of z_ν's of the lemma. Since from (15,2) and (10,7) we have

$$T_L > L > \log^{\frac{1}{30}} T > \log^{\frac{1}{60}} T > |\varrho_0|,$$

$\varrho = \varrho_0$ occurs in Z, i. e., the condition

$$\max_\nu |z_\nu| \geqq 1$$

is certainly fulfilled. We choose

(16,1) $$m = K_0.$$

The n of the lemma is obviously the number $N(T_L)$ of the zeros of $\zeta(s)$ in

$$0 < \sigma < 1, \qquad |t| \leqq T_L;$$

owing to $L < T_L < L + 1$ we may choose for the N of the lemma any upper bound for $N(L+1)$ of the zeros in

$$0 < \sigma < 1, \qquad |t| \leqq L + 1.$$

Riemann-Mangoldt's theorem gives in the weak form[1]

$$N(L+1) < c_{11} L \log L;$$

hence from (10,7)

$$N(L+1) < \frac{c_{11}}{10} \log^{\frac{1}{10}} T \log\log T,$$

i. e., we may choose

(16,2) $$N = \frac{c_{11}}{10} \log^{\frac{1}{10}} T \log\log T.$$

Further we have

$$m = K_0 = 10\frac{\log T}{\log\log T} > 28\frac{c_{11}}{10} \log^{\frac{1}{10}} T \log\log T = 28\, N$$

if

$$100 \log^{0\cdot 9} T > 28\, c_{11} (\log\log T)^2$$

which is true for $T > c_{12}$. Hence our lemma is applicable to Z; the only question which remains is whether the value so obtained for $(k+1)$ fulfils the restriction (10,9) or not. Denoting the value of $(k+1)$ given by lemma by $(k^* + 1)$, we have for $T > c_{13}$

$$K_0 = m \leqq k^* + 1 \leqq m + N = K_0 + \frac{c_{11}}{10} \log^{\frac{1}{10}} T \log\log T <$$

(16,3) $$< K_0 + \log^{\frac{1}{10}} T\, (\log\log T)^2 = K_0 + N_0,$$

i. e., (10,9) is indeed satisfied by choosing $k = k^*$. Hence for $T > c_{14}$

$$|Z| > \frac{100}{c_{11}^2} \frac{1}{\log^{\frac{1}{5}} T (\log\log T)^2} \cdot \left(\frac{c_{11}}{10^2 e^{27}} \frac{\log^{\frac{1}{10}} T (\log\log T)^2}{\log T}\right)^{\frac{c_{11}}{10} \log^{\frac{1}{10}} T \log\log T} >$$

$$> e^{-c_{15} \log^{\frac{1}{10}} T (\log\log T)^2}$$

or from (15,3)

$$|J_1| > \frac{T^{\beta_0}}{|\varrho_0|^{10 \frac{\log\log T}{\log T}}} e^{-\left(c_{15} + \frac{1}{5}\right) \log^{\frac{1}{10}} T (\log\log T)^2}$$

and from (14,8)

$$\frac{1}{(k^*)!} \int_1^{\xi} |\Delta(x)| \left|\frac{d}{dx} \log^{k^*} \frac{\xi}{x}\right| dx > \frac{T^{\beta_0}}{|\varrho_0|^{10 \frac{\log\log T}{\log T}}} e^{-\left(c_{15} + \frac{1}{5}\right) \log^{\frac{1}{10}} T (\log\log T)^2} -$$

$$(16,4) \qquad - (c_{10} + 1)\, e^{41 \frac{\log T \log\log\log T}{\log\log T}}.$$

If $c_{16}(\varrho_0)$ is determined so that for $T > c_{16}(\varrho_0)$

$$(16,5) \qquad (c_{10} + 1)\, e^{41 \frac{\log T \log\log\log T}{\log\log T}} < \frac{1}{2} \frac{\sqrt{T}}{|\varrho_0|^{10 \frac{\log T}{\log\log T}}} e^{-\left(c_{15} + \frac{1}{5}\right) \log^{\frac{1}{10}} T (\log\log T)^2},$$

then we obtain for all such large T-values

$$B = \frac{1}{(k^*)!} \int_1^{\xi} |\Delta(x)| \left|\frac{d}{dx} \log^{k^*} \frac{\xi}{x}\right| dx > \frac{1}{2} \frac{T^{\beta}}{|\varrho_0|^{10 \frac{\log T}{\log\log T}}} \cdot e^{-\left(c_{15} + \frac{1}{5}\right) \log^{\frac{1}{10}} T (\log\log T)^2}.$$

(16,6)

17. To complete the proof of our theorem we have only to find an upper estimation for the integral on the left side. Introducing the notation

$$\max_{1 \leq x \leq T} |\Delta(x)| = M,$$

first we have from (16,3) and (10,6)

$$\xi = L^{k^*+1} \leq L^{K_0 + N_0} \leq T,$$

$$\max_{1 \leq x \leq \xi} |\Delta(x)| \leq M,$$

i. e.,

$$B = \frac{1}{(k^*)!} \int_1^{\xi} |\Delta(x)| \left|\frac{d}{dx} \log^{k^*} \frac{\xi}{x}\right| dx \leq -\frac{M}{(k^*)!} \int_1^{\xi} \frac{d}{dx} \left(\log^{k^*} \frac{\xi}{x}\right) dx <$$

$$< \frac{M}{(k^*)!} \log^{k^*} T < M \left(\frac{e \log T}{k^*}\right)^{k^*}.$$

But

$$\frac{1}{k^*} < \frac{2}{k^*+1} \leqq \frac{2}{K_0} = \frac{1}{5}\,\frac{\log\log T}{\log T},$$

$$k^* < 2\,K_0 = 20\,\frac{\log T}{\log\log T},$$

i. e.,

$$B < M\left(\frac{e}{5}\log\log T\right)^{20\frac{\log T}{\log\log T}} < M\,e^{40\frac{\log T\log\log\log T}{\log\log T}}$$

or from (16,6) and (10,2)

$$M > \frac{1}{2}\,\frac{T^{\beta_0}}{|\varrho_0|^{10\frac{\log T}{\log\log T}}}\,e^{-\left(c_{15}+\frac{1}{5}\right)\log^{\frac{1}{10}}T\,(\log\log T)^3-40\frac{\log T\log\log\log T}{\log\log T}} >$$

$$> \frac{T^{\beta_0}}{|\varrho_0|^{10\frac{\log T}{\log\log T}}}\,e^{-c_{17}\frac{\log T\log\log\log T}{\log\log T}}$$

with

$$c_{17} = c_{15} + 41. \qquad Q.\,e.\,d.$$

(Received 31 March 1950)

ОБ ОСТАТОЧНОМ ЧЛЕНЕ ФОРМУЛЫ ПРОСТЫХ ЧИСЕЛ. I

ПАВЕЛ ТУРАН (Будапешт)

(Резюме)

Пусть $\Lambda(n)$ — хорошо известный символ Диришлэ, $\psi(x) = \sum_{n<x} \Lambda(n)$ и $\Delta(x) = \psi(x) - x$. Фрагмен, Э. Шмидт, Ландау, Литтльвуд, Пигэм и другие авторы занимались оценками $\Delta(x)$ и подобных отношений типов Ω и $\Omega\pm$. Известные до сих пор результаты отличаются общим свойством зависимости от конфигурации всех нетривиальных корней функции $\zeta(s)$ и поэтому, а даже и вследствие своих методов, они не дают *явных* оценок. Для устранения этих затруднений, Литтльвуд предлагает задать вопрос возможно-ли дать оценку Ω для $\Delta(x)$, зависящего лишь от одного (произвольного) нетривиального корня функции $\zeta(s)$. В настоящей статье, автор дает результат такого типа, как новое применение своего метода, изложенного на с'езде чехо-словацкого и польского математических обществ в Праге в 1949 г. Здесь доказана следующая теорема:

Если $\varrho_0 = \beta_0 + i\gamma_0$ — произвольный нетривиальный корень $\zeta(s)$, то можно дать *явно* две числовые постоянные c_0 и c_1, далее $c_2 = c_2(\varrho_0)$, так что для $T > \max(c_0, c_2(\varrho_0))$ мы имеем

$$\max_{1\leq x\leq T}(\Delta(x)) > T^{\beta_0}\,|\varrho_0|^{-\frac{10\log T}{\log\log T}}\,e^{-\frac{c_1\log T\log\log\log T}{\log\log T}}.$$

CONTRIBUTIONS TO THE REDUCTION THEORY OF THE DECISION PROBLEM

First paper

Prefix $(x_1)\,(x_2)\,(Ex_3)\ldots(Ex_{n-1})\,(x_n)$, a single binary predicate

By

LÁSZLÓ KALMÁR (Szeged), corresponding member of the Academy

1. In the reduction theory of the decision problem of the first order predicate calculus, the intention is to find as simple cases of this problem as possible, which are equivalent to the decision problem itself. On the one hand, the *number of arguments* of the predicate variables contained in the formula to be examined has been reduced: indeed, LÖWENHEIM[1] has proved that the special case of the decision problem dealing with binary[2] formulae is equivalent to the general problem. In addition, I have reduced the *number of the predicate variables* appearing in the formula to one.[3] On the other hand, it has been attempted to reduce the decision problem to the case if the formula to be examined as to being satisfiable has a *special type of prefix*, as simple

[1] L. LÖWENHEIM, Über Möglichkeiten im Relativkalkül, *Mathematische Annalen*, **76** (1915), pp. 447—470, especially theorem 6, pp. 463—470.

[2] A further reduction, viz. the reduction to the case of unary formulae, is impossible, for this case of the decision problem has been solved by LÖWENHEIM (loc. cit., especially theorem 4, pp. 459—462) by an algorithm which is readily to be recognized as a recursive one (in the sense of ALONZO CHURCH, An unsolvable problem of elementary number theory, *American Journal of Mathematics*, **58** (1936), pp. 345—363 and S. C. KLEENE, General recursive functions of natural numbers, *Mathematische Annalen*, **112** (1936), pp. 727—742), while the decision problem itself has been proved to be unsolvable by any such algorithm by ALONZO CHURCH, A note on the Entscheidungsproblem, *The Journal of Symbolic Logic*, **1** (1936), pp. 40—41 and 101—102.

[3] LÁSZLÓ KALMÁR, Zurückführung des Entscheidungsproblems auf den Fall von Formeln mit einer einzigen, binären, Funktionsvariablen, *Compositio Mathematica*, **4** (1936), pp. 137—144; see also Zum Entscheidungsproblem der mathematischen Logik, *Verhandlungen des internationalen Mathematikerkongresses Zürich* 1932, vol. 2, pp. 337—338. Previously, HERBRAND has reduced the same number to three; see JACQUES HERBRAND, Sur le problème fondamental de la logique mathématique, *Sprawozdania z posiedzeń Towarzystwa Naukowego Warszawskiego*, wydział III, **24** (1931), pp. 12—56, especially pp. 39—41.

as possible.[4] However, the simplicity of a prefix can be judged from different points of view. The Skolem prefix[5]

$$(x_1) \ldots (x_m) (Ex_{m+1}) \ldots (Ex_n)$$

is the simplest one as to the *number of alternations* of universal and existential quantifiers.[6] As to its subcases, the Gödel prefix[7]

(1) $$(x_1)(x_2)(x_3)(Ex_4) \ldots (Ex_n)$$

contains a *minimal number of universal quantifiers*,[8] while the Pepis prefix[9]

(2) $$(x_1) \ldots (x_{n-1})(Ex_n)$$

keeps the *number of existential quantifiers* as small as possible. The "non-Skolemian" Pepis prefix[10]

(3) $$(x_1)(x_2)(Ex_3)(x_4) \ldots (x_n)$$

has the advantage over (2) of reducing the *order of the existential quantifier*

[4] In the first line, see the papers quoted in the footnotes [5], [7], [8] (papers of PEPIS), [9], [10] and [12].

[5] See TH. SKOLEM, Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze nebst einem Theorem über dichte Mengen, *Videnskapsselkapets Skrifter*, Kristiania, mat.-naturv. klasse, 1920, no. 4, pp. 1—36, especially pp. 4—6.

[6] Indeed, for the case of a prefix without alternation of universal and existential quantifiers at all, (i. e. of the form $(x_1) \ldots (x_n)$ or $(Ex_1) \ldots (Ex_n)$) the satisfiability problem has been solved by BERNAYS and SCHÖNFINKEL by a recursive algorithm; the same holds for the other case of a prefix with a single alternation, viz. of the form $(Ex_1) \ldots (Ex_m)(x_{m+1}) \ldots (x_n)$. See P. BERNAYS and M. SCHÖNFINKEL, Zum Entscheidungsproblem der mathematischen Logik, *Mathematische Annalen*, **99** (1928), pp. 342—372, especially pp. 359—360.

[7] See K. GÖDEL, Zum Entscheidungsproblem des logischen Funktionenkalküls, *Monatshefte für Mathematik und Physik*, **40** (1933), pp. 433—443.

[8] Indeed, for the case of a Skolem prefix with two universal quantifiers, the decision problem has been solved by a recursive algorithm independently by GÖDEL, KALMÁR and SCHÜTTE; see K. GÖDEL, Ein Spezialfall des Entscheidungsproblems der theoretischen Logik, *Ergebnisse eines mathematischen Kolloquiums*, **2** (1932), pp. 27—28; LÁSZLÓ KALMÁR, Über die Erfüllbarkeit derjenigen Zählausdrücke, welche in der Normalform zwei benachbarte Allzeichen enthalten, *Mathematische Annalen*, **108** (1933), pp. 466—484; KURT SCHÜTTE, Untersuchungen zum Entscheidungsproblem der mathematischen Logik, *ibid.*, **109** (1934), pp. 572—603 and Über die Erfüllbarkeit einer Klasse von logischen Formeln, *ibid.*, **110** (1935), pp. 161—194. If we do not insist on the Skolem form, and if we allow the fixed predicate $x = y$ besides predicate variables, a further reduction of the number of universal quantifiers, viz. to two, is possible; see JÓZEF PEPIS, Beiträge zur Reduktionstheorie des logischen Entscheidungsproblems, *Acta Scientiarum Mathematicarum*, **8** (1936—37), pp. 7—41, especially theorems 38 and 39, pp. 37—41; and Untersuchungen über das Entscheidungsproblem des mathematischen Logik, *Fundamenta Mathematicae*, **30** (1938), pp. 257—348, especially theorem 33, p. 319. A still further reduction is impossible, owing to a recursive solution of the satisfiability problem of formulae containing a single universal quantifier; see WILHELM ACKERMANN, Über die Erfüllbarkeit gewisser Zählausdrücke, *Mathematische Annalen*, **100** (1928), pp. 638—649.

[9] See JÓZEF PEPIS, Ein Verfahren der mathematischen Logik, *The Journal of Symbolic Logic*, **3** (1938), pp. 61—76 and Untersuchungen etc. (loc. cit.), theorems 36 and 39, pp. 339—340.

[10] See JÓZEF PEPIS, Ein Verfahren etc. (loc. cit.) and Untersuchungen etc. (loc. cit.), theorems 37 and 40, pp. 339—340.

from $n-1$ to 2; here, the order of an existential quantifier is defined as the number of universal quantifiers preceding it.[11] The Ackermann prefix[12]

$$(4) \qquad (Ex_1)(x_2)(Ex_3)(x_4)\ldots(x_n)$$

is — though containing one more existential quantifier — still more advantageous in this respect, the order of the quantifiers being zero and one, respectively.

As to the *simultaneous* reduction of the decision problem in both directions, it has been proved in three preceding papers[13] that the decision problem is equivalent to those of its special cases in which the formula to be examined as to being satisfiable, besides containing no predicate variable but a single binary one, has a prefix of the form (4), resp. (1), resp. (2) or (3).

Now, the question arises as to whether the same is true for the prefixes

$$(5) \qquad (x_1)(x_2)(Ex_3)\ldots(Ex_{n-1})(x_n)$$

and

$$(6) \qquad (x_1)(Ex_2)\ldots(Ex_{n-2})(x_{n-1})(x_n),$$

similar to (1) but, though non-Skolemian, more advantageous as to the order of the existential quantifiers. In this paper, I shall answer this question as to the prefix (5) by proving the

THEOREM. *To any given formula of the first order predicate calculus it is possible to construct another, containing no other predicate variable but a single binary one, having a prefix of the form* (5), *and equivalent, in respect of being or being not satisfiable, to the given formula.*

As to the prefix (6), I shall return to the question in a subsequent paper.[14]

2. To prove our theorem, first, we construct to any given first order formula **A** another, say **B**, which, besides having a prefix of the form (5) and a matrix containing a single, binary, predicate variable, can be satisfied if **A** can; then, we prove that conversely, if **B** can be satisfied, the same holds or **A** too.

[11] By an existential quantifier the existence of an individual depending on the individual variables bound by universal quantifiers preceding it is formalized; i. e. that of a descriptive function over the set of individuals, the number of the arguments of which equals the order of the existential quantifier in question.

[12] See WILHELM ACKERMANN, Beiträge zum Entscheidungsproblem der mathematischen Logik, *Mathematische Annalen*, **112** (1936), pp. 419—432.

[13] LÁSZLÓ KALMÁR, On the reduction of the decision problem, first paper, Ackermann prefix, a single binary predicate, *The Journal of Symbolic Logic*, **4** (1939), pp. 1—9 LÁSZLÓ KALMÁR and JÁNOS SURÁNYI, On the reduction etc., second paper, Gödel prefix, a single binary predicate, *ibid.*, **12** (1947), pp. 65—73; third paper, Pepis prefix, a single binary predicate, *ibid.*, **15** (1950) pp. 161—173. I shall quote the first and the second one of these papers as "Reduction I" and "Reduction II."

[14] We observe that neither the theorem of the present paper is an immediate consequence of that of Reduction II, nor conversely; and the analogous theorem concerning the prefix (6) is, in the same sense, independent of the theorem of the present paper as well as of that of Reduction II.

By the theorem of Reduction II, we may suppose, **A** has the prefix (1) and a matrix

(7) $$\mathbf{M} = \mathbf{M}(F;\ x_1, \ldots, x_n)$$

containing the binary predicate variable F but no other; thus, **M** is a truth-function of the n^2 arguments $F(x_\mu, x_\nu)$, $\mu, \nu = 1, \ldots, n$.

Suppose **A** can be satisfied on a set I by a binary predicate Φ defined over I; i. e., for some descriptive funtions $\xi_4(x_1, x_2, x_3), \ldots, \xi_n(x_1, x_2, x_3)$ defined over I, we have

(8) $$\mathbf{M}(\Phi;\ x_1, x_2, x_3, \xi_4(x_1, x_2, x_3), \ldots, \xi_n(x_1, x_2, x_3))$$

for arbitrary $x_1, x_2, x_3 \in I$. The formula **B** has to give an alternative formalization of the same fact. In order to replace the three initial universal quantifiers by two, we adjoin to I the ordered pairs of elements of I; thus, "for arbitrary $x_1, x_2, x_3 \in I$" we can express as "for arbitrary $x \in I^2$ and $y \in I$," I^2 denoting the set of ordered pairs formed of elements of I, x standing for the pair (x_1, x_2) and y for x_3. In order to assure adequacy of this new expression to the old one, **B** has to formalize also, for any $x, y \in I$, the existence of a $p \in I^2$ with x as first and y as second component. The predicates "$x \in I$," "$x \in I^2$," "x is the first component of p," "x is the second component of p" are to be expressed, beside "$\Phi(x, y)$," by means of the single predicate Ψ satisfying **B**. For this purpose we adjoin, besides the elements of I^2, the predicate Φ to I, and, on the new domain of individuals[15] $J = I + I^2 + \{\Phi\}$ resulting thus, we define the binary predicate Ψ so that

(i) $\Psi(x, \Phi)$ holds if and only if $x \in I$;

(ii) $\Psi(\Phi, p)$ holds if and only if $p = (x, y) \in I^2$ $(x, y \in I)$ and $\Phi(x, y)$ holds;

(iii) for $x \in I$, $p \in I^2$, $\Psi(x, p)$ holds if and only if x is the first component of p,

(iv) for $x \in I$, $p \in I^2$, $\Psi(p, x)$ holds if and only if x is the second component of p,

finally, in order to distinguish between the cases[16] $x = \Phi$ and $x \in I^2$,

(v) $\Psi(x, x)$ holds always excepting for $x = \Phi$,

so that[17]

(vi) $\Psi(x, x)\ \overline{\Psi}(x, \Phi)$ holds if and only if $x \in I^2$.

[15] We suppose tacitly that I and I^2 have no element in common; i. e., no element of I is of the form (x, y) with $x, y \in I$. We can secure this by replacing I, if necessary, by an appropriate set of the same cardinal number. Of course, Φ does not belong to I or I^2.

[16] By (i), we have $\overline{\Psi}(x, \Phi)$ in both cases. — Of course, instead of the predicate Ψ defined above we could use another predicate $\Psi(x, y)$ for which $\Psi(x, x)$ holds if and only if $x = \Phi$ (and, $\overline{\Psi}(x, \Phi)$ holds if and only if $x \in I$); this predicate would be more analogous to those chosen in Reduction I and Reduction II. However, in our case the definition adopted here seems more natural.

[17] For simplicity, we omit the conjunction sign, or we replace it by a dot.

We see at once that conditions (i)—(vi) are consistent[18] and they define the predicate $\Psi(x, y)$ over J except for $x, y \in I$, $x \neq y$, and for $x, y \in I^2$, $x \neq y$. Defining $\Psi(x, y)$ in these cases to be true,[19] say, we can tabulate the definition of Ψ as follows

	$y = \Phi$	$y \in I$	$y = (y_1, y_2) \in I^2$
$x = \Phi$	false	false	$\Phi(y_1, y_2)$
$x \in I$	true	true	$x = y_1$
$x = (x_1, x_2) \in I^2$	false	$x_2 = y$	true

Now, we could formalize the existence of a pair having given components as follows[20]

$$(9) \qquad (x)\,(y)\,(Ep)\,(\Psi(x, \Phi)\,\Psi(y, \Phi) \to \Psi(p, p)\,\overline{\Psi}(p, \Phi)\,\Psi(x, p)\,\Psi(p, y)).$$

However, similarly as in Reduction II, we have to ensure that the predicate $\Psi(\Phi, p)$ in (ii) depends on x and y only. In our case this can be done much simpler than in Reduction II,[21] for we have an universal quantifier at disposal which can be placed after the existential one;[22] thus, we can extend (9) to

$$(10) \qquad (x)\,(y)\,(Ep)\,(z)\,\{\Psi(x, \Phi)\,\Psi(y, \Phi) \to \Psi(p, p)\,\overline{\Psi}(p, \Phi)\,\Psi(x, p)\cdot$$
$$\cdot\,\Psi(p, y)\,[\Psi(z, z)\,\overline{\Psi}(z, \Phi)\,\Psi(x, z)\,\Psi(z, y) \to (\Psi(\Phi, z) \sim \Psi(\Phi, p$$

3. Now, let $x = (x_1, x_2) \in I^2$ $(x_1, x_2 \in I)$ and $y \in I$. Let $x_3 = y$ and

$$x_\mu = \xi_\mu(x_1, x_2, x_3) \quad \text{for } \mu = 4, \ldots, n,$$
$$p_{\mu\nu} = (x_\mu, x_\nu) \quad \text{for } \mu, \nu = 1, \ldots, n.$$

Then we have, by (i),

$$(11) \qquad \Psi(x_\mu, \Phi) \text{ for } \mu = 4, \ldots, n;$$

[18] Indeed, (vi) is a consequence of (i) and (v) (see footnote [15]); and (i)—(v) determine the truth-value of Ψ at different places except $\Psi(\Phi, \Phi)$, which is determined to be false by each of (i), (ii) or (v).

[19] $\Psi(x, y)$ has to be true, on account of (v), for $x = y \in I$ as well as for $x = y \quad I^2$. We have chosen $\Psi(x, y)$ to be true for $x, y \in I$, $x \neq y$ and for $x, y \in I^2$, $x \neq y$ merely to simplify the following table. For a similar reason, in (ii), we have defined $\Psi(\Phi, x)$ to be false for $x \in I$.

[20] Here and in the following formulae (up to the end of 5), J has to be taken as the range of the quantifiers.

[21] In Reduction II, we had to introduce two representatives of each pair to the same effect.

[22] Here, the situation is analogous to that in my paper: László Kalmár, Zur Reduktion des Entscheidungsproblems, *Norsk Matematisk Tidsskrift*, 19 (1937), pp. 121—130, especially pp. 127—129.

by (vi),

(12) $$\Psi(p_{\mu\nu}, p_{\mu\nu})\,\overline{\Psi}(p_{\mu\nu}, \Phi) \quad \text{for } \mu, \nu = 1, \ldots, n;$$

by (iii) and (iv),

(13) $$\Psi(y, p_{3\nu})\,\Psi(p_{\nu 3}, y) \quad \text{for } \nu = 1, \ldots, n$$

and

(14) $$\Psi(x_\mu, p_{\mu\nu})\,\Psi(p_{\mu\nu}, x_\nu) \quad \text{for } \mu = 4, \ldots, n \text{ and } \nu = 1, \ldots, n;$$

further, for arbitrary $z \in J$, by (i), (iii) and (iv),

(15) $$\Psi(z, \Phi)\,\Psi(z, x) \rightarrow \Psi(z, p_{1\nu})\,\Psi(p_{\nu 1}, z) \quad \text{for } \nu = 1, \ldots, n$$

and

(16) $$\Psi(z, \Phi)\,\Psi(x, z) \rightarrow \Psi(z\ p_{2\nu})\,\Psi(p_{\nu 2}, z) \quad \text{for } \nu = 1, \ldots, n.$$

Finally, denoting by $\mathbf{M}^* = \mathbf{M}^*(G: u, p_{11}, p_{12}, \ldots, p_{nn})$ the matrix which results from (7) by replacing $F(x_\mu, x_\nu)$ throughout by $G(u, p_{\mu\nu})$ for $\mu, \nu = 1, \ldots, n$, we have by (8) and (ii),

(17) $$\mathbf{M}^*(\Psi; \Phi, p_{11}, p_{12}, \ldots, p_{nn}).$$

Summarizing our results, we have[22] by (vi), (i), (11), (12), (13), (14), (15), (16) and (17),

(18) $$(x)(y)(Ex_4)\ldots(Ex_n)\,Ep_{11})(Ep_{12})\ldots(Ep_{nn})(z)\,\Big\{\Psi(x, x)\cdot$$
$$\overline{\Psi}(x, \Phi)\,\Psi(y, \Phi) \rightarrow \prod_{\mu=4}^{n}\Psi(x_\mu, \Phi)\prod_{\mu,\nu=1}^{n}\big(\Psi(p_{\mu\nu}, p_{\mu\nu})\,\overline{\Psi}(p_{\mu\nu}, \Phi)\big)\cdot$$
$$\cdot\prod_{\nu=1}^{n}\Big[\big(\Psi(z, \Phi)\,\Psi(z, x) \rightarrow \Psi(z, p_{1\nu})\,\Psi(p_{\nu 1}, z)\big)\big(\Psi(z, \Phi)\,\Psi(x, z) \rightarrow$$
$$\rightarrow \Psi(z, p_{2\nu})\,\Psi(p_{\nu 2}, z)\big)\,\Psi(y, p_{3\nu})\,\Psi(p_{\nu 3}, y)\cdot$$
$$\cdot\prod_{\mu=4}^{n}\big(\Psi(x_\mu, p_{\mu\nu})\,\Psi(p_{\nu\mu}, x_\mu)\big)\Big]\,\mathbf{M}^*(\Psi; \Phi, p_{11}, p_{12}, \ldots, p_{nn})\Big\}.$$

4. Now, we have to express the existence of an element Φ of J with the properties (10) and (18). To this effect, the most natural way would be to form the conjunction of (10) and (18), to replace Φ throughout by an individual variable u and to place the quantifier (Eu) at the beginning of the formula resulting thus. This way would lead, choosing an appropriate prenex form for the conjunction of (10) and (18), to the prefix

$$(Eu)(x)(y)(Ep)(Ex_4)\ldots(Ex_n)(Ep_{11})(Ep_{12})\ldots(Ep_{nn})(z),$$

not of the form (5). However, we can avoid this inconvenience similarly as in Reduction II, viz. by weakening (10) and (18) to the proposition that for any $x, y \in J$, there is a $u \in J$ having the same property as Φ in (10) and (18), further, securing independence of this u from x and y by postulating that it has, in addition, the property $\overline{\Psi}(u, u)$ characterizing Φ, and,

[22] For simplicity, we abbreviate conjunctions of analogous terms with a Π-sign.

that any two elements x and y of J having this property can be replaced. each by the other, in either argument of Ψ, no matter what stands in the other argument, thus, they can be replaced, each by the other, in any formula containing no predicate but Ψ. The last postulate can be formalized as follows:

$$(19)\quad (x)(y)(z)\left[\overline{\Psi}(x,x)\,\overline{\Psi}(y,y) \to \big(\Psi(x,z) \sim \Psi(y,z)\big)\big(\Psi(z,x) \sim \Psi(z,y)\big)\right],$$

and the stated modifications on the conjunction of (10) and (18) lead to[24]

$$\begin{aligned}(20)\quad & (x)(y)(Eu)(Ep)(Ex_4)\ldots(Ex_n)(Ep_{11})(Ep_{12})\ldots(Ep_{nn})(z)\Big\{\Big[\Psi(x,u)\cdot\\ &\cdot\Psi(y,u) \to \Psi(p,p)\,\Psi(p,u)\,\Psi(x,p)\,\Psi(p,y)\cdot\\ &\cdot\big[\Psi(z,z)\,\overline{\Psi}(z,u)\,\Psi(x,z)\,\Psi(z,y) \to \big(\Psi(u,z)\sim\Psi(u,p)\big)\big]\Big]\cdot\\ &\cdot\Big[\Psi(x,x)\,\overline{\Psi}(x,u)\,\Psi(y,u) \to \prod_{u=4}^{n}\Psi(x_\mu,u)\prod_{\mu,\nu=1}^{n}\big(\Psi(p_{\mu\nu},p_{\mu\nu})\,\overline{\Psi}(p_{u\nu},u)\big)\cdot\\ &\cdot\prod_{\nu=1}^{n}\big[\big(\Psi(z,u)\,\Psi(z,x) \to \Psi(z,p_{1\nu})\,\Psi(p_{\nu 1},z)\big)\cdot\\ &\cdot\big(\Psi(z,u)\,\Psi(x,z) \to \Psi(z,p_{2\nu})\,\Psi(p_{\nu 2},z)\,\Psi(y,p_{3\nu})\,\Psi(p_{\nu 3},y)\cdot\\ &\cdot\prod_{\mu=4}^{n}\big(\Psi(x_u,p_{\mu\nu})\,\Psi(p_{\nu\mu},x_u)\big)\big]\,\mathbf{M}^*(\Psi: u, p_{11}, p_{12},\ldots, p_{nn})\Big]\,\overline{\Psi}(u,u)\Big\}.\end{aligned}$$

5. Forming the conjunction of (19) and (20), carrying it to a prenex form and changing Ψ throughout into a predicate variable G, we get the required formula[25] **B**, viz.

$$\begin{aligned} & (x)(y)(Eu)(Ep)(Ex_4)\ldots(Ex_n)(Ep_{11})(Ep_{12})\ldots(Ep_{nn})(z)\Big\{\Big[\overline{G}(x,x)\cdot\\ &\cdot\overline{G}(y,y) \to \big(G(x,z)\sim G(y,z)\big)\big(G(z,x)\sim G(z,y)\big)\Big]\cdot\\ &\cdot\Big[G(x,u)\,G(y,u) \to G(p,p)\,\overline{G}(p,u)\,G(x,p)\,G(p,y)\cdot\\ &\cdot\big[G(z,z)\,\overline{G}(z,u)\,G(x,z)\,G(z,y) \to \big(G(u,z)\sim G(u,p)\big)\big]\Big]\cdot\\ &\cdot\Big[G(x,x)\,G(x,u)\,G(y,u) \to \prod_{\mu=4}^{n}G(x_u,u)\prod_{\nu,\mu=1}^{n}\big(G(p_{\mu\nu},p_{u\nu})\cdot\\ &\cdot\overline{G}(p_{u\nu},u)\big)\prod_{\nu=1}^{n}\big[\big(G(z,u)\,G(z,x)\,G(z,p_{1\nu})\,G(p_{\nu 1},z)\big)\cdot\end{aligned}$$

[24] Formally, we can deduce (20) from (10), (18) and $\Psi(\Phi,\Phi)$ by omitting the quantifiers $(x)\,(y)$ (on account of the rule "$(x)\,F(x)$ implies $F(x)$"), by forming the conjunction, by changing Φ into an individual variable u bound by an existential quantifier (Eu) (on account of the rule "$F(\Phi)$ implies $(Eu)\,F(u)$") and by placing the quantifiers $(x)\,(y)$ at the beginning of the formula (on account of the rule "$F(x)$ implies $(x)\,F(x)$").

[25] By a slight modification of the above argument, we could spare two existential quantifiers, viz. by omitting the quantifiers, (Ep) and (Ep_{12}) and replacing p by p_{13} and p_{12} by x in the matrix. However, these modifications would complicate our formulae.

$$\cdot \left(G(z, u)\, G(x, z) \rightarrow G(z, p_{2\nu})\, \bar{G}(p_{\nu 2}, z)\right) G(y, p_{3\nu})\, G(p_{\nu 3}, y) \cdot$$

$$\cdot \prod_{\mu=4}^{n} \left(G(x_\mu, p_{\mu\nu})\, G(p_{\nu\mu}, x_\mu)\right)\Big] \quad \mathbf{M}^*(G;\, u, p_{11}, p_{12}, \ldots, p_{nn})\Big]\, \bar{G}(u, u)\Big\} .$$

We see at once, **B** has a prefix of the form (5) and can be satisfied, by what precedes, if **A** can.

6. To prove the converse of the last property of **B**, suppose, **B** can be satisfied on a (non empty) set J' by a predicate Ψ to be substituted for G. Then, (19) and (20) hold, J' being taken as range of the quantifiers. By (20), there are descriptive functions $u = \varphi(x, y)$, $p = \pi(x, y)$, $x_\mu = \xi_\mu(x, y)$ $(\mu = 4, \ldots, n)$, $p_{\mu\nu} = \pi_{\mu\nu}(x, y)$ $\quad(\mu, \nu = 1, \ldots, n)$ defined over J' such that for any $x, y \in J'$ for which $\Psi\left(x, \varphi(x, y)\right)$ and $\Psi\left(y, \varphi(x, y)\right)$ hold, we have

(9′) $$\Psi\left(\pi(x, y), \pi(x, y)\right) \bar{\Psi}\left(\pi(x, y), \varphi(x, y)\right),$$

(9‴) $$\Psi\left(x, \pi(x, y)\right) \Psi\left(\pi(x, y), y\right);$$

further, for any $z \in J'$ for which $\Psi(z, z)\, \Psi\left(z, \varphi(x, y)\right) \Psi(x, z)\; \Psi(z, y)$ holds, we have

(10′) $$\Psi\left(\varphi(x, y), z\right) \sim \Psi\left(\varphi(x, y), \pi(x, y)\right);$$

moreover, for any $x, y \in J'$ for which $\Psi(x, x)\, \overline{\Psi}\left(x, \varphi(x, y)\right)$ and $\Psi\left(y, \varphi(x, y)\right)$ hold, we have

(11′) $$\Psi\left(\xi_\mu(x, y), \varphi(x, y)\right) \text{ for } \mu = 4, \ldots, n,$$

(12′) $$\Psi\left(\pi_{\mu\nu}(x, y), \pi_{\mu\nu}(x, y)\right) \overline{\Psi}\left(\pi_{\mu\nu}(x, y), \varphi(x, y)\right) \text{ for } \mu, \nu = 1, \ldots, n,$$

(13′) $$\Psi\left(y, \pi_{3\nu}(x, y)\right) \Psi\left(\pi_{\nu 3}(x, y), y\right) \text{ for } \nu = 1, \ldots, n,$$

(14′) $$\Psi\left(\xi_\mu(x, y), \pi_{\mu\nu}(x, y)\right) \Psi\left(\pi_{\nu\mu}(x, y), \xi_\mu(x, y)\right)$$
$$\text{for } \mu = 4, \ldots, n \text{ and } \nu = 1, \ldots, n,$$

(17′) $$\mathbf{M}^*\left(\Psi;\, \varphi(x, y), \pi_{11}(x, y), \pi_{12}(x, y), \ldots, \pi_{nn}(x, y)\right),$$

and, for any $z \in J'$,

(15′) $$\Psi\left(z, \varphi(x, y)\right) \Psi(z, x) \rightarrow \Psi\left(z, \pi_{1\nu}(x, y)\right) \Psi\left(\pi_{\nu 1}(x, y), z\right)$$

and

(16′) $$\Psi\left(z, \varphi(x, y)\right) \Psi(x, z) \rightarrow \Psi\left(z, \pi_{2\nu}(x, y)\right) \Psi\left(\pi_{\nu 2}(x, y), z\right) \text{ for } \nu = 1, \ldots, n;$$

finally, we have for any $x, y \in J'$ whatever,

(21) $$\overline{\Psi}\left(\varphi(x, y), \varphi(x, y)\right).$$

7. Let a be any fixed element of J' and let $\varphi = \varphi(a, a)$. Then, we get from (21)

$$\overline{\Psi}(\varphi, \varphi);$$

hence, considering (21), we obtain from (19) for arbitrary $x, y, z \in J'$

$$\Psi\left(\varphi(x, y), z\right) \sim \Psi(\varphi, z)$$

and

$$\Psi(z, \varphi(x, y)) \sim \Psi(z, \varphi);$$

therefore, we may replace $\varphi(x, y)$ by φ in (9′), (10′), (11′), (12′), (17′), (15′), (16′) as well as in the conditions $\Psi(x, \varphi(x, y))$, $\Psi(y, \varphi(x, y))$, $\overline{\Psi}(z, \varphi(x, y))$ and $\overline{\Psi}(x, \varphi(x, y))$. We shall quote the formulae modified thus as (9″), (10″), (11″), (12″), (17″), (15″) and (16″).

8. We define I' as the subset of J' defined by the condition $\Psi(x, \varphi)$. On account of (11″), we see that I' is not empty. We shall prove that **A** can be satisfied on I'.

Let x_1, x_2, x_3 be arbitrary elements of I'. Let

$$x = \pi(x_1, x_2),$$

$$y = x_3,$$

$$x_\mu = \xi_\mu(x, y) = \xi_\mu(\pi(x_1, x_2), x_3) \text{ for } \mu = 4, \ldots, n,$$

and

$$p_{\mu\nu} = \pi_{\mu\nu}(x, y) = \pi_{\mu\nu}(\pi(x_1, x_2), x_3) \text{ for } \mu, \nu = 1, \ldots, n.$$

By (9″), we have (taking x_1 for x and x_2 for y)

$$\Psi(x, x)\overline{\Psi}(x, \varphi),$$

further, on account of $y \in I'$,

$$\Psi(y, \varphi),$$

so that the conditions of (11″), (12″), (13′), (14′), (17″), (15″) and (16″) are fulfilled. (11″) shows that $x_4, \ldots, x_n \in I'$; from (12″) we take

(22) $$\Psi(p_{\mu\nu}, p_{\mu\nu})\overline{\Psi}(p_{\mu\nu}, \varphi) \text{ for } \mu, \nu = 1, \ldots, n.$$

Now, we show

(23) $$\Psi(x_\mu, p_{\mu\nu})\Psi(p_{\nu\mu}, x_\mu) \text{ for } \mu, \nu = 1, \ldots, n.$$

For $\mu = 4, \ldots, n$ and $\nu = 1, \ldots, n$, (23) is identical to (14′); (13′) shows that (23) remains true for $\mu = 3$; $\nu = 1, \ldots, n$ too. By (9‴) we have (taking x_1 for x and x_2 for y)

(24) $$\Psi(x_1, x)$$

and

(25) $$\Psi(x, x_2).$$

From (15″), $z = x_1$, we obtain, on account of $x_1 \in I'$ and (24),

$$\Psi(x_1, p_{1\nu})\Psi(p_{\nu 1}, x_1) \text{ for } \nu = 1, \ldots, n;$$

similarly, from (16″), $z = x_2$, we obtain

$$\Psi(x_2, p_{2\nu})\Psi(p_{\nu 2}, x_2) \text{ for } \nu = 1, \ldots, n,$$

since $x_2 \in I'$ and (25) holds. That is, (23) holds for $\mu = 1,2$; $\nu = 1, \ldots, n$ too.

Now, (23) can be written as follows:

$$\Psi(x_\mu, p_{\mu\nu})\, \Psi(p_{\mu\nu}, x_\nu) \text{ for } \mu, \nu = 1, \ldots, n;$$

this shows, together with (22), that the conditions of (10″) are fulfilled by taking x_μ, x_ν and $p_{\mu\nu}$ for x, y and z, respectively; therefore, we have

$$\Psi(\varphi, p_{\mu\nu}) \sim \Psi(\varphi, \pi(x_\mu, x_\nu)) \text{ for } \mu, \nu = 1, \ldots, n.$$

Thus, we may replace in (17″), i. e., in

$$\mathbf{M}^*\ (\Psi; \varphi, p_{11}, p_{12}, \ldots, p_{nn})$$

the proposition $\Psi(\varphi, p_{\mu\nu})$ by $\Psi(\varphi, \pi(x_\mu, x_\nu))$ for $\mu, \nu = 1, \ldots, n$. On account of the definition of $\mathbf{M}^*$, this means that for the predicate Φ defined for arbitrary $x, y \in I'$ by

$$\Phi(x, y) = \Psi(\varphi, \pi(x, y)),$$

we have

$$\mathbf{M}\ (\Phi; x_1, \ldots, x_n);$$

x_1, x_2 and x_3 being arbitrary elements of I', this shows that **A** is satisfied by Φ on I'. Thus the equivalence of **A** and **B** has been proved, and so our theorem holds.

The Bolyai Institute, University of Szeged.

(Received 16 April 1950)

ВКЛАДЫ В ТЕОРИЮ ПРИВЕДЕНИЯ ПРОБЛЕМЫ РАЗРЕШИМОСТИ

Первая статья. Префикс $(x_1)\ (x_2)\ (Ex_3) \ldots (Ex_{n-1})\ (x_n)$, единственный, двухпеременнный предикат.

ЛАСЛО КАЛЬМАР (Сегед)

(Резюме)

Автор, в совместной статье с Яношом Шурани (Journal of Symbolic Logic, том 12-ый), доказал, что проблема разрешимости математической логики может быть приведена к вопросу выполнимости таких логических формул с префиксом Геделя $(x_1\ x_2\ x_3\ Ex_4) \ldots Ex_n$, которые содержат только одну, двухпеременную логическую переменную функцию. Исходя из этого результата, автор в настоящем статье доказывает аналогичную теорему, относящуюся к префиксу, упоминаемому в подзаголовке, которая постольку преимущесвеннее, поскольку экзистенциальные кванторы опережаются не тремя, а только двумя общими кванторами, следовательно они выражают существование не трехпеременных, а двухпеременных математических функций. Метод доказательства состоит в замене x_1 и x_2 одной упорядоченной парой $(x_1, x_2$. То, что каждому x_1 и x_2 принадлежит одна и только одна упорядоченная пара, у которой первый компонент x_1, а второй компонент x_2, удается выразить без увеличения числа логических функций, с помощью единственного общего квантора, ставящегося после экзистенциальных кванторов.

ZUR THEORIE DER FAKTORISIERBAREN GRUPPEN. I

Von

L. RÉDEI (Szeged), korrespondierendem Mitglied der Akademie

§ 1. Einleitung

Stets bezeichnen G, Γ, $\mathfrak{G}$ drei Gruppen mit

$$\mathfrak{G} = G\Gamma, \tag{1}$$

unter welcher Gleichung wir verstehen, daß die Produkte $A\mathsf{A}$ ($A \in G$, $\mathsf{A} \in \Gamma$) alle verschiedenen Elemente von $\mathfrak{G}$ sind. Wir nennen dann $\mathfrak{G}$ eine *faktorisierbare Gruppe* mit den komplementären Faktoren G, Γ. Gleich bemerken wir, daß die Reihenfolge von G, Γ unwesentlich ist, denn (1) und $\mathfrak{G} = \Gamma G$ bedingen einander gegenseitig. (Das verleiht der Theorie der faktorisierbaren Gruppen eine Art Dualität, s. unten.) Die faktorisierbaren Gruppen haben ZAPPA [11][1] und SZÉP [5]—[10] voneinander unabhängig untersucht, gewisse ihrer Resultate finden sich teils ergänzt auch in einer neueren Arbeit von mir [2]. Ein Teil der SZÉPschen Arbeiten betrachtet den allgemeineren Fall, daß in (1) die gesagten Produkte $A\mathsf{A}$ die Elemente von $\mathfrak{G}$ auch mehrmals darstellen dürfen, von diesem ebenfalls sehr interessanten Fall nehmen wir hier Abstand. ZAPPA hat nicht nur die Eigenschaften der faktorisierbaren Gruppen untersucht, sondern er hat auch das folgende wichtige *Umkehrproblem* gestellt und gelöst: Für zwei gegebene Gruppen G, Γ sind alle Gruppen $\mathfrak{G}$ mit der Eigenschaft (1) zu bestimmen. Dabei dürfen G, Γ (als abstrakte Gruppen) ganz beliebig sein, damit aber das Problem lösbar ist, müssen wir G, Γ natürlich so realisieren, daß sie dasselbe Einselement aber keine weiteren gemeinsamen Elemente haben. Das nehmen wir ohne Einschränkung der Allgemeinheit stets an und bezeichnen mit $A, B, C, \ldots, P, \ldots, X$ und $\mathsf{A}, \mathsf{B}, \mathsf{C}, \ldots \Pi, \ldots \mathsf{X}$ Elemente von G bzw. Γ. Das Einselement bezeichnen wir mit E bzw. E, so daß dann nach unserer Annahme $E = \mathsf{E}$ gilt. (Es wird von Nutzen sein für das gemeinsame Einselement beide Bezeichnungen E, E beizubehalten.) Der Satz von ZAPPA lautet so:[2]

[1] Eckige Klammern verweisen auf das Literaturverzeichnis angefügt am Ende der Arbeit.

[2] Vgl. die Arbeit [2], in der ich weniger einfach (A, A) für $A\mathsf{A}$ geschrieben habe, diese Abweichung ist unwesentlich, verleiht aber dem Satz von ZAPPA eine überaus durchsichtige Form. Der Satz läßt sich aus den Gruppenpostulaten mit wenig Rechnung ableiten (ohne Kenntnis der Arbeiten [11] oder [2]).

Gilt (1) für die Gruppen G, Γ, $\mathfrak{G}$, so gibt es zwei eindeutig bestimmte Funktionen (in „Operatorschreibweise")

(2) $$A^{\mathsf{A}}\ (\in G), \qquad \mathsf{A}^{A}\ (\in \Gamma)$$
mit

(3) $$\mathsf{A}A = A^{\mathsf{A}}\mathsf{A}^{A},$$
es gilt auch

(4) $$A^{\mathsf{E}} = A, \qquad E^{\mathsf{A}} = E, \qquad \mathsf{A}^{E} = \mathsf{A}, \qquad \mathsf{E}^{A} = \mathsf{E}$$

(5) $$A^{\mathsf{BC}} = (A^{\mathsf{C}})^{\mathsf{B}}, \qquad \mathsf{A}^{BC} = (\mathsf{A}^{B})^{C},$$

(6) $$(AB)^{\mathsf{C}} = A^{\mathsf{C}}B^{\mathsf{C}^{A}}, \qquad (\mathsf{AB})^{C} = \mathsf{A}^{C^{\mathsf{B}}}\mathsf{B}^{C}.$$

Umgekehrt, wenn die Gruppen G, Γ vorgegeben sind und man eine Lösung (2) von (4), (5), (6) bestimmt, so erzeugen die Elemente von G und Γ bei der Annahme (3) eine Gruppe $\mathfrak{G}$ [3] *mit der Eigenschaft (1), alle letzteren Gruppen gewinnt man auf diesem Wege.*

Zur Ergänzung bemerken wir, daß das Multiplikationsgesetz in $\mathfrak{G}$ wegen (3) so lautet:

(7) $$A\mathsf{A}B\mathsf{B} = AB^{\mathsf{A}}\mathsf{A}^{B}\mathsf{B}.$$

Die erste Hälfte des Satzes von ZAPPA hat auch SZÉP gewonnen, außerdem hat er zuerst aus den Bedingungsgleichungen (4), (5), (6) weitere interessante Schlüsse gezogen. Aus diesen Gründen nennen wir jede Gruppe $\mathfrak{G}$ in (1) ein ZAPPA—SZÉPsches Produkt von G und Γ.

Ich muß auf die auch schon durch ZAPPA bemerkte Analogie mit der SCHREIERschen Erweiterungstheorie hinweisen, in der es sich bekanntlich um die Bestimmung der Gruppen $\mathfrak{G}$ mit $\mathfrak{G}/G \cong \Gamma$ handelt. Wie ich gezeigt habe [2], läßt sich die SCHREIERsche Erweiterungstheorie ähnlich begründen wie die Theorie von ZAPPA—SZÉP, ich fand sogar, daß sich beide als Spezialfälle in eine allgemeinere Theorie einbetten lassen. Aus diesem Grunde nenne ich die SCHREIERsche Erweiterung von G mit Γ gerne auch das SCHREIERsche Produkt von G und Γ. Mir scheinen beide Produkte (das SCHREIERsche und das ZAPPA—SZÉPsche) gleich wichtige Angelegenheiten der Gruppentheorie zu sein.

Ein ZAPPA—SZÉPsches Produkt $\mathfrak{G} = G\Gamma$ enthält unter Umständen Normalteiler, ein Teil der Untersuchungen von SZÉP gibt eben solche Fälle an, es kann aber auch einfach sein, z. B. gehört die Ikosaedergruppe auch hierzu (vgl. [2]). Wenn eben G oder Γ ein Normalteiler von $\mathfrak{G} = G\Gamma$ ist, dann und nur dann stimmen beide Produkte überein und zwar ist das eben die sogenannte zerfallende Erweiterung. Wir werden in dieser Arbeit auch Beispielen

[3] Das ist so zu verstehen, daß die Elemente von $\mathfrak{G}$ alle endlichen Produkte aus den Elementen von G, Γ sind, die sich aber wegen (3) auf die Form $A\mathsf{A}$ bringen lassen.

begegnen, wo $\mathfrak{G} = G\Gamma$ wohl nicht einfach aber kein SCHREIERsches Produkt von G und Γ oder von Γ und G ist.

Weiter wird uns hier nur das oben formulierte Umkehrproblem von ZAPPA interessieren, wo nämlich G, Γ vorgegeben sind und man alle $\mathfrak{G} = G\Gamma$ sucht, in dieser Richtung machte man bisher keine Untersuchungen. Der Satz von ZAPPA bringt dieses Problem gar nicht zum Abschluß, sondern führt es auf die Angabe aller Lösungen (2) der Funktionalgleichungen (4), (5), (6) zurück. Die Lösungen zu bestimmen ist im allgemeinen eine sehr schwierige Aufgabe, überhaupt scheint uns die Theorie von ZAPPA—SZÉP schwieriger und vielgestalteter zu sein als die von SCHREIER.

Ohne Zweifel ist eine der primitivsten Aufgaben der ZAPPA—SZÉPschen Theorie, das ZAPPA—SZÉPsche Produkt $G\Gamma$ von zwei zyklischen Gruppen G, Γ zu bestimmen. Auch schon dieser Fall, der uns hier von einigen vorbereitenden allgemeineren Untersuchungen abgesehen allein beschäftigen wird, bietet ernste Schwierigkeiten. Den Spezialfall, daß die eine der zyklischen Gruppen G, Γ unendlich, die andere endlich ist, konnten wir vollständig erledigen, für den Fall aber, daß G, Γ beide unendlich sind, haben wir nur einige Teilresultate bekommen, die vollständige Lösung ist eine sehr interessante und schwierige Aufgabe. In der Mitteilung II dieser Arbeit wollen wir den Fall betrachten, daß G, Γ beide endlich sind, dieser Fall ist überraschend reich an Lösungen. (Für die SCHREIERsche Theorie ist der Fall von zyklischen G, Γ bekanntlich sehr leicht.)

Wir wollen hier einiges über den allgemeinen Fall bemerken. Kraft (2) erscheint im Satz von ZAPPA jedes von G, Γ als ein Operatorenbereich für das andere. Und zwar nennen wir A und A in einer (symbolischen) Potenz A^{A} das *Grundelement* bzw. den *Operator* (für A^{A} vertauschen sich die Rollen). Die Gleichungen (5) besagen mit der üblichen Redensart, daß G und Γ je ein multiplikativer Operatorenbereich für das andere ist (und zwar das erste ein Rechts-, das zweite ein Linksoperatorenbereich). Aus (5_1), (4_1) sieht man sofort, daß die Abbildung $\mathrm{A} \to \mathrm{A}^{\mathsf{A}}$ (bei festem A) eine Permutation von G ist, und zwar hat die Gleichung $\mathrm{A}^{\mathsf{A}} = \mathrm{B}$ bei festen A, B die einzige Lösung $\mathrm{A} = \mathrm{B}^{\mathsf{A}^{-1}}$, ähnlich ist $\mathsf{A} \to \mathsf{A}^{\mathrm{A}}$ eine Permutation von Γ. Während man es aber in der Gruppentheorie meistens mit solchen Operationen A^{A} zu tun hat, für die $(\mathrm{A}\,\mathrm{B})^{\mathsf{A}} = \mathrm{A}^{\mathsf{A}}\,\mathrm{B}^{\mathsf{A}}$ gilt („linearer“ oder „distributiver Operator“), so daß dann die gesagte Permutation $\mathrm{A} \to \mathrm{A}^{\mathsf{A}}$ ein Automorphismus von G ist, in unserem Falle gelten die viel komplizierteren Bedingungen (6), eben darin liegt die Schwierigkeit der Theorie von ZAPPA—SZÉP.

Des näheren bemerken wir hierzu folgendes. Bezeichnen wir für einen Augenblick mit $\mathrm{A}_1, \mathrm{A}_2, \ldots$ bzw. $\mathsf{A}_1, \mathsf{A}_2, \ldots$ je ein System von Erzeugenden von G und Γ. Man würde sich denken, daß das Funktionenpaar (2) durch die Bedingungsgleichungen (4), (5), (6) und durch die „Anfangswerte“ $\mathrm{A}_i^{\mathsf{A}_k}$,

$\mathsf{A}_k^{\mathsf{A}_i}$ eindeutig bestimmt ist (wie im Falle linearer Operatoren), aber wir werden die Falschheit hiervon mit einem Beispiel zeigen.

Wir bemerken, daß im ZAPPA—SZÉPschen Produkt $G\Gamma$ oft eine gewisse *Asymmetrieerscheinung* auftritt, womit wir folgendes meinen. Wohl sind (4), (5), (6) symmetrisch in G und Γ (die völlige Symmetrie könnte man so herstellen, daß man z. B. in Γ von AB auf die umgekehrte Multiplikation BA übergeht), trotzdem, wenn G und Γ isomorph sind, kommt es vor, daß in einer Lösung (2) die eine der Funktionen α^{A}, A^{α} viel komplizierter aussieht als die andere, man könnte in solchen Fällen sagen, daß im ZAPPA—SZÉPschen Produkt $G\Gamma$ der eine Faktor den anderen überwältigt. Auch hierfür werden wir ein interessantes Beispiel sehen.

Es muß bemerkt werden, daß die triviale Lösung $\alpha^{\mathsf{A}} = \alpha$, $\mathsf{A}^{\alpha} = \mathsf{A}$ stets vorliegt, dieser entspricht das direkte Produkt von G und Γ.

§ 2. Allgemeines

Bezeichne H eine beliebig vorgelegte Gruppe. Für gewöhnlich versteht man unter einem Komplex K (von H) eine Teilmenge von H, oft kommt es aber auch vor, daß man dabei den Elementen des Komplexes eine Multiplizität zukommen läßt. Das wollen wir tun, so daß wir einen Komplex in der Form

(8) $$K = a_1\gamma_1, a_2\gamma_2, \ldots$$

annehmen, wobei $\gamma_1, \gamma_2, \ldots$ alle verschiedenen Elemente von H sind und jedes a_i gleich $0, 1, 2, \ldots$ oder ∞ ist. (Ist H von der endlichen Ordnung n, so ist (8) mit dem Glied $a_n\gamma_n$ abzubrechen.) Kommt ein $a_i = \infty$ vor, so nennen wir K *uneigentlich*. Im Falle $a_1 = a_2 = \ldots = 0$ setzen wir $K = 0$. Die $a_i\gamma$ nennen wir die Glieder von K, zwei Komplexe sind gleich zu betrachten, wenn ihre Glieder (ungeachtet der Reihenfolge) übereinstimmen. Selbst a_i nennen wir die Multiplizität oder den Koeffizienten von γ_i in K. Die Glieder $a_i\gamma_i$ mit $a_i = 0$ dürfen wir in (8) streichen. Dann und nur dann, wenn $a_i \neq 0$ ist, sagen wir, daß γ_i ein *Element von* K ist, hierfür verwenden wir die mengentheoretische Schreibweise $\gamma_i \in K$. Ist

(9) $$L = b_1\gamma_1, b_2\gamma_2, \ldots$$

ein weiterer Komplex, so definieren wir die Summe beider Komplexe durch

$$K + L = (a_1 + b_1)\gamma_1, (a_2 + b_2)\gamma_2, \ldots$$

und entsprechend soll die Summe beliebig (eventuell unendlich) vieler Komplexe definiert werden. Selbstverständlich is dabei eine Summe $a + b + \ldots$ $(a, b, \ldots = 0, 1, \ldots, \infty)$ gleich ∞ zu setzen, wenn entweder unendlich viele nichtverschwindende Summanden vorkommen oder mindestens ein

Summand ∞ vorkommt. Die Addition von Komplexen ist offenbar kommutativ und assoziativ. Nunmehr dürfen wir (8) auch als $K = a_1 \gamma_1 + a_2 \gamma_2 + \cdots$ schreiben. Wir werden diese „additive Schreibweise der Komplexe" bevorzugen. Ferner definieren wir das Produkt $K L$ als die Summe aller $a_i b_j \gamma_i \gamma_j$, dabei ist ein Produkt $a b$ $(a, b = 0, 1, \ldots, \infty)$ gleich 0 zu setzen, wenn $a = 0$ oder $b = 0$ gilt, und es ist $ab = \infty$ zu setzen, wenn weder $a = 0$ noch $b = 0$ gilt und $a = \infty$ oder $b = \infty$ ist. Die Multiplikation der Komplexe ist assoziativ und (gegenüber der Addition) beiderseits distributiv. (Auch das Produkt unendlich vieler Komplexe könnte man entsprechend definieren, aber wir nehmen davon Abstand.) Wohl unterschieden von der Potenz K^i $(i = = 1, 2, \ldots)$ bezeichne $K^{(i)}$ $(i = \pm 1, \pm 2, \ldots)$ denjenigen Komplex, den man aus (8) bei Ersetzung von jedem γ_k durch γ_k^i gewinnt. Insbesondere nennen wir $K^{(-1)}$ den zu K inversen Komplex. Endlich definieren wir das Skalarprodukt rK $(r = 0, 1, \ldots, \infty)$ durch $rK = r\, a_1 \gamma_1, r\, a_2 \gamma_2, \ldots$ Eine (endliche oder unendliche) Summe

$$r_1 K_1 + r_2 K_2 + \ldots \qquad (r_i = 0, 1, \ldots, \infty) \tag{10}$$

nennen wir eine Linearkombination der Komplexe $K_1, K_2, \ldots$. Sind diese nicht uneigentlich, von 0 verschieden und paarweise elementenfremd, so sind die r_i durch den Wert von (10) eindeutig bestimmt.

Wir verabreden uns noch, daß wir für ein Glied $a_i \gamma_i$ von K mit $a_i = 1$ einfach γ_i schreiben. Durch diese Vereinbarung haben wir die gewöhnlichen Komplexe unter die obigen verallgemeinerten Komplexe eingeordnet, sie lassen sich dann auch in der Form $\alpha + \beta + \ldots$ schreiben (mit verschiedenen $\alpha, \beta, \ldots$). Umgekehrt ist K in (8) dann und nur dann ein gewöhnlicher Komplex, wenn alle a_i gleich 0 oder 1 sind. Da insbesondere eine Untergruppe U von H stets ein gewöhnlicher Komplex ist, deshalb führen wir im Bereich unserer verallgemeinerten Komplexe die kurze (aber nicht ganz korrekte) Redensart „ein (verallgemeinerter) Komplex K ist eine Gruppe" ein und verstehen darunter, daß K ein gewöhnlicher Komplex ist und seine Elemente eine Gruppe bilden. Diese Redensart werden wir auch dann verwenden, wenn wir K in additiver Form schreiben, ein Mißverständnis wird daraus nicht entstehen[4].

Von hier an betrachten wir die anfangs eingeführten Gruppen G, Γ und ein Funktionenpaar (2) mit den Eigenschaften (4), (5), (6). Ein (verallgemei-

[4] Die oben eingeführten verallgemeinerten Komplexe und die Rechnung mit ihnen bilden ein wichtiges Hilfsmittel in der Theorie von ZAPPA—SZÉP, das uns aus den nachfolgenden Entwicklungen klar wird. Wir bemerken folgendes. Würde man in unseren Komplexen nur endlich viele Elemente und nur endliche aber auch negative Multiplizitätszahlen zulassen, so erhält man den wohlbekannten Begriff des Gruppenringes (über dem Ring der ganzen Zahlen). Dieser Gruppenring spielt eine unentbehrliche Rolle im Beweis des Satzes von HAJÓS [1] über endliche ABELsche Gruppen, in dem es sich ebenfalls um ein gewisses (sehr schwieriges) Faktorisationsproblem handelt. Vgl. RÉDEI [3] und SZELE [4].

nerter) Komplex soll nachher stets einen solchen von G oder von Γ bedeuten. Ist $K = a_1 A_1 + a_2 A_2 + \ldots$ ein Komplex von G, so setzen wir $K^{\mathsf{A}} = a_1 A_1^{\mathsf{A}} + a_2 A_2^{\mathsf{A}} + \ldots$. Gilt $K^{\mathsf{A}} = K$ für alle A ($\in \Gamma$), so nennen wir den Komplex K *operatorinvariant.* Im Spezialfall, wo $K = A$ ein Element von G ist, sprechen wir dann über ein *operatorinvariantes Element,* wofür wir auch *Fixelement* sagen. Das Einselement E ist stets ein Fixelement, allgemeiner folgt aus (6_1) sofort der

Hilfssatz 1. *Die Fixelemente von G bilden stets eine Untergruppe von G.*

Die vorangeschickten Definitionen und selbst Hifssatz 1 gelten wörtlich auch nach Vertauschung von G mit Γ. Diese Vertauschung nennen wir *Dualisation.* Auch der noch folgende Inhalt dieses Paragraphen läßt sich dualisieren, worüber wir unten noch näheres sagen werden.

Von einem beliebigen Element A ($\in G$) ausgehend üben wir auf ihm alle Operatoren A ($\in \Gamma$) aus. Die verschiedenen der erhaltenen Elemente A^{A} (A fest, A beliebig variabel) bilden einen (gewöhnlichen) Komplex von G, den wir mit $T = T(A)$ bezeichnen und die Transitivitätsklasse von A nennen. Stets gilt dann $A \in T(A)$. Wir beweisen den

Hilfssatz 2. *Die Transitivitätsklassen $T(A)$ sind operatorinvariant und die verschiedenen unter ihnen bilden eine Klasseneinteilung von G in paarweise fremde Klassen.*

Da nämlich $A \to A^{\mathsf{A}}$ eine Permutation von G ist, so folgt hieraus, daß für einen gewöhnlichen Komplex K von G stets auch K^{A} ein gewöhnlicher Komplex von G ist. Insbesondere ist also jedes $(T(A))^{\mathsf{A}}$ ein gewöhnlicher Komplex. Gilt nun $B \in T(A)$ d. h. $B = A^{\mathsf{B}}$ mit einem passenden B, so folgt aus (5_1) $B^{\mathsf{A}} = A^{\mathsf{AB}}$. Dies ergibt $(T(A))^{\mathsf{A}} \leq T(A)$ für alle A. Dann gilt auch $(T(A))^{\mathsf{A}^{-1}} \leq T(A)$, also $T(A) \leq T((A))^{\mathsf{A}}$, beide ergeben $(T(A))^{\mathsf{A}} = T(A)$, folglich ist $T(A)$ operatorinvariant. Haben $T(A)$, $T(B)$ ein gemeinsames Element $A^{\mathsf{A}} = B^{\mathsf{B}}$, so folgt aus (5_1) $A^{\mathsf{C}} = B^{\mathsf{C}\mathsf{A}^{-1}\mathsf{B}}$ also $T(A) \leq T(B)$. Aus Symmetriegründen folgt $T(A) = T(B)$, womit Hilfssatz 2 bewiesen ist.

Nach der Definition sind die Transitivitätsklassen mit nur einem Element eben die Fixelemente.

Hilfssatz 3. *Die Linearkombinationen der Transitivitätsklassen sind die sämtlichen operatorinvarianten Komplexe von G.*

Daß nämlich eine Linearkombination der $T(A)$ operatorinvariant ist, ist wegen Hilfssatz 2 klar. Bezeichne umgekehrt K einen operatorinvarianten Komplex von G und A ein beliebiges Element von G. Dann kommen A und A^{A} ($\mathsf{A} \in \Gamma$) mit derselben Multiplizität in K vor, und dies bedeutet, daß für jede Transitivitätsklasse $T(A)$ sämtliche Elemente von $T(A)$ die gleiche Multiplizität in K haben. Dies und Hilfssatz 2 liefern den Beweis von Hilfssatz 3.

Hilfssatz 4. *Bezeichne T und T je eine Transitivitätsklasse von $\mathring{G}$ bzw. Γ. Dann gelten die Formeln:*

$$(11)\qquad (\mathsf{A}\, T)^{\mathsf{A}} = \mathsf{A}^{\mathsf{A}} T, \quad (\mathsf{T}\,\mathsf{A})^{\mathsf{A}} = \mathsf{T}\,\mathsf{A}^{\mathsf{A}} \quad (\mathsf{A} \in G,\ \mathsf{A} \in \Gamma).$$

(Auch diese Formeln sind dual, in ihnen kommt die sich in (5), (6) offenbarende Asymmetrie in Vorschein. Es ist TA nicht mit $\mathsf{T(A)}$ zu verwechseln!)

Aus Dualitätsgründen genügt es (11_1) zu beweisen. Man lasse B in (6_1) alle Elemente von T durchlaufen und addiere die so erhaltenen Gleichungen, wegen $T = T^{\mathsf{A}}$ (A beliebig) entsteht dann

$$(\mathsf{A}\, T)^{\mathsf{C}} = \mathsf{A}^{\mathsf{C}} T.$$

Dies stimmt mit (11_1) überein, womit wir Hilfssatz 4 bewiesen haben.

Wir beweisen jetzt die folgendenden zwei Hilfsätze:

Haupthilfssatz 1. *Für zwei beliebige Transitivitätsklassen T_1, T_2 von G gilt*

$$(12)\qquad (T_1 T_2)^{\mathsf{A}} = T_1 T_2 \qquad (\mathsf{A} \in \Gamma).$$

Dies bedeutet, daß das Produkt von zwei Transitivitätsklassen von G stets operatorinvariant, d. h. nach Hilfssatz 3 eine Linearkombination der Transitivitätsklassen von G ist. In Formeln:

$$(13)\qquad T_i T_k = \sum_l c_{i,k,l} T_l \quad (c_{i,k,l} = 0, 1, \ldots, \infty),$$

wobei T_1, T_2, ... alle Transitivitätsklassen von G bezeichnen.

Haupthilfssatz 2. *Für jede Transitivitätsklasse T von G ist auch der inverse Komplex $T^{(-1)}$ eine Transitivitätsklasse.* Gilt $T = T^{(-1)}$, so nennen wir die Transitivitätsklasse *selbstinvers*[5].

[5] Für den Augenblick nennen wir allgemeiner jede Klasseneinteilung T_1, T_2, ... einer Gruppe G regulär, wenn für diese beide Haupthilfssätze 1,2 gelten und das Einselement für sich eine Klasse bildet. Die regulären Klasseneinteilungen spielen in der Gruppentheorie eine wohlbekannte wichtige Rolle. Und zwar bezeichne jetzt Γ (anders als im Text) eine Automorphismengruppe von G (deren Elemente wir weiter auch auf die bisherige Art bezeichnen), mit anderen Worten sollen jetzt die [(4), (5), (6) ähnlichen aber einfacheren] Gleichungen $\mathrm{A}^{\mathsf{E}} = \mathrm{A}$, $\mathrm{E}^{\mathsf{A}} = \mathrm{E}$, $\mathrm{A}^{\mathsf{BC}} = (\mathrm{A}^{\mathsf{C}})^{\mathsf{B}}$, $(\mathrm{AB})^{\mathsf{C}} = \mathrm{A}^{\mathsf{C}}\mathrm{B}^{\mathsf{C}}$ gelten. Definiert man dann die Klasse $T = T(\mathrm{A})$ wieder als die Menge aller A^{A} ($\mathsf{A} \in \Gamma$), so entsteht offenbar eine reguläre Klasseneinteilung von G. (Das allerbekannteste Beispiel bilden die natürlichen Klassen von G, so nenne ich die Klassen der konjugierten Elemente. Dies entspricht dem Fall, daß Γ die Gruppe der inneren Automorphismen von G ist. Ist Γ die volle Automorphismengruppe von G, so nenne ich die entsprechenden Klassen charakteristisch.) Es braucht nicht besonders betont zu werden, daß wegen der Haupthilfssätze 1,2 für die Theorie von Zappa—Szép die Untersuchung der regulären Klasseneinteilungen einer Gruppe eine wichtige Aufgabe ist. Um Mißverständnisse zu vermeiden bemerken wir, daß in einer regulären Klasseneinteilung einer Gruppe sehr wohl Elemente verschiedener Ordnung in eine Klasse kommen können (insbesondere bei den natürlichen und den charakteristischen Klassen ist das nie der Fall). Z. B. ist

1; (14) (23); (12) (34), (142), (124), (143), (134); (13) (24), (123), (132), (234), (243)

eine Einteilung der alternierenden Permutationsgruppe von vier Ziffern in vier reguläre Klassen, wie man das leicht nachprüfen kann, dabei finden sich in einer Klasse Elemente zweiter und dritter Ordnung.

Bemerkung. Selbstverständlich gilt das Duale der Haupthilfssätze 1,2 auch.

Für den Haupthilfssatz 1 genügt es (12) zu beweisen. Nun folgt aber (12) sofort aus (11_1), wenn man in diesem für A die Elemente einer Transitivitätsklasse von G einsetzt und summiert, wobei man die Operatorinvarianz der Transitivitätsklassen zu berücksichtigen hat.

Den Haupthilfssatz 2 beweisen wir so. Gilt $\mathsf{A} \in T^{(-1)}$, so gilt $\mathrm{E} \in \mathrm{A}\,T$, und dann folgt hieraus nach (11_1) $\mathrm{E} \in \mathrm{A}^{\mathsf{A}}\,T$, also auch $\mathrm{A}^{\mathsf{A}} \in T^{(-1)}$. Das bisherige besagt $(T^{(-1)})^{\mathsf{A}} < T^{(-1)}$. Da dies auch für A^{-1} statt A gilt, leitet sich aus beiden $(T^{(-1)})^{\mathsf{A}} = T^{(-1)}$ ab. Hiernach ist $T^{(-1)}$ operatorinvariant, und so gilt nach Hilfssatz 3 eine Gleichung $T^{(-1)} = T_1 + T_2 + \ldots$ mit gewissen Transitivitätsklassen T_i. Dann besteht $T = T_1^{(-1)} + T_2^{(-1)} + \ldots$. Da aber jedes Glied der rechten Seite nach dem schon bewiesenen ebenfalls eine Linearkombination von Transitivitätsklassen ist, so folgt aus der Bemerkung über die Eindeutigkeit von (10) mit Notwendigkeit $T = T_1^{(-1)}$ also $T^{(-1)} = T_1$. Haupthilfssatz 2 haben wir bewiesen.

Als weitere Vorbereitung bezeichnen wir die Permutation $\mathrm{A} \to \mathrm{A}^{\mathsf{A}}$ mit $[\mathsf{A}]$. Ähnlich bezeichne $[\mathrm{A}]$ die Permutation $\mathsf{A} \to \mathsf{A}^{\mathrm{A}}$. Nach (5) bilden diese Permutationen (nämlich die verschiedenen unter ihnen) je eine Gruppe, die wir mit $[\Gamma]$ bzw. $[G]$ bezeichnen, und es gelten die Homomorphien:

$$G \sim [G], \quad \Gamma \sim [\Gamma]. \tag{14}$$

Und zwar werden diese Homomorphien bzw. durch die Abbildung

$$\mathrm{A} \to [\mathrm{A}], \quad \mathsf{A} \to [\mathsf{A}^{-1}] \tag{15}$$

vermittelt. Ein Element $\mathrm{A}\,(\in G)$ oder $\mathsf{A}\,(\in \Gamma)$ nennen wir einen *identischen Operator*, wenn $[\mathrm{A}]$ bzw. $[\mathsf{A}]$ die identische Permutation (von Γ bzw. G) ist. Die identischen Operatoren in G bzw. Γ bilden je eine Untergruppe G^0, Γ^0, diese sind die beiden Kerne der Homomorphien (14), folglich gelten die Isomorphien

$$G/G^0 \approx [G], \qquad \Gamma/\Gamma^0 \approx [\Gamma]. \tag{16}$$

Aus (5_1) sieht man sofort, daß gewiß $\mathrm{A}^{\mathsf{A}} = \mathrm{A}^{\mathsf{B}}$ gilt, wenn A, B dieselbe Klasse nach Γ^0 repräsentieren. Läßt man also A ein volles Repräsentantensystem der Klassen von Γ mod Γ^0 durchlaufen, so durchläuft A^{A} auch schon alle Elemente der Transitivitätsklasse $T(\mathrm{A})$ (einige eventuell mehrmals). Wenn also Γ/Γ^0 endlich ist, so sind auch alle Transitivitätsklassen endlich.

Aus den Definitionen und aus (6_1) entstehen sofort:

Hilfssatz 5. *Ist C ein Fixelement von Γ, so gilt*

$$(\mathrm{A}\,\mathrm{B})^{\mathsf{C}} = \mathrm{A}^{\mathsf{C}}\,\mathrm{B}^{\mathsf{C}}.$$

(Kurz: Die Fixelemente von Γ bewirken als Operatoren lauter Automorphismen von G.)

Hilfssatz 6. *Sind die Elemente* A, B ($\in G$) *identische Operatoren* (*für* Γ), *so gilt*

$$(\mathrm{A}\,\mathrm{B})^{\mathsf{C}} = \mathrm{A}^{\mathsf{C}}\,\mathrm{B}^{\mathsf{C}}.$$

(Kurz: Bei Anwendung eines Operators in Γ erleidet die Gruppe G^0 der identischen Operatoren in G einen Automorphismus.)

Wir bemerken noch die aus (6) folgenden Formeln: Ist C bzw. C je ein Fixelement von G bzw. Γ, so gilt

$$(17)\qquad (\mathrm{A}\,\mathsf{C})^{\mathsf{A}} = \mathrm{A}^{\mathsf{A}}\mathsf{C}, \qquad (\mathsf{C}\,\mathsf{A})^{\mathrm{A}} = \mathsf{C}\,\mathsf{A}^{\mathrm{A}}.$$

(Beide Hilfssätze 5, 6 lassen sich dualisieren. Die Formeln (17) sind zueinander „asymmetrisch-dual".)

§ 3. Der Fall zyklischer Gruppen G, Γ

Wie üblich bezeichnen wir mit $\}\ \{$ die durch die eingeklammerten Elemente erzeugte Gruppe. Von jetzt an betrachten wir in dieser Arbeit weiter nur den Fall, daß die Gruppen G, Γ beide zyklisch sind und setzen

$$(18)\qquad G = \}\mathrm{P}\{, \quad \Gamma = \}\pi\{,$$

wobei P, π je ein erzeugendes Element ist. Mit $O(x)$ bezeichnen wir die Ordnung von x, wobei x eine Gruppe oder ein Gruppenelement ist.

Für zyklische (sogar allgemeiner für Abelsche) Gruppen G, Γ verschwindet die sich in den Formeln (5), (6) offenbarende Asymmetrie, und zwar gilt jetzt

$$(19)\qquad \mathrm{A}^{\mathsf{BC}} = (\mathrm{A}^{\mathsf{B}})^{\mathsf{C}} = (\mathrm{A}^{\mathsf{C}})^{\mathsf{B}}, \qquad \mathsf{A}^{\mathrm{BC}} = (\mathsf{A}^{\mathrm{B}})^{\mathrm{C}} = (\mathsf{A}^{\mathrm{C}})^{\mathrm{B}},$$

$$(20)\qquad (\mathrm{A}\,\mathrm{B})^{\mathsf{C}} = \mathrm{A}^{\mathsf{C}^{\mathrm{B}}}\,\mathrm{B}^{\mathsf{C}} = \mathrm{A}^{\mathsf{C}}\,\mathrm{B}^{\mathsf{C}^{\mathrm{A}}}, \qquad (\mathsf{A}\,\mathsf{B})^{\mathrm{C}} = \mathsf{A}^{\mathrm{C}^{\mathsf{B}}}\,\mathsf{B}^{\mathrm{C}} = \mathsf{A}^{\mathrm{C}}\,\mathsf{B}^{\mathrm{C}^{\mathsf{A}}}.$$

Aus (14) folgt, daß die Permutationsgruppen $[G]$, $[\Gamma]$ auch zyklisch sind, offenbar sind dabei [P], $[\pi]$ erzeugende Elemente. Zerlegt man die Permutation [P] (von Γ) in paarweise fremde Zyklen, so fallen diese mit den Transitivitätsklassen von Γ zusammen, das Duale hiervon gilt auch.

Wie bemerkt, wollen wir in dieser Arbeit den Fall außer Acht lassen, daß G, Γ beide endlich sind, deshalb nehmen wir im folgenden

$$(21)\qquad O(G) = \infty$$

an. Die hier zu beweisenden Hilfssätze gelten dann, ob Γ endlich oder unendlich ist.

Hilfssatz 7. *Ist T ($\neq$ E) eine endliche Transitivitätsklasse von G mit minimaler Elementzahl, so ist auch jedes $T^{(i)}$ eine solche.*

Wegen $T^{(-i)} = (T^{(-1)})^{(i)}$ und Haupthilfssatz 2 genügt es die Behauptung für $i > 0$ zu beweisen. Bezeiche a die größte ganze Zahl mit $\mathrm{P}^a \in T$. Aus dem

Polynomialsatz folgt, daß die außerhalb von $T^{(i)}$ liegenden Elemente von $T^{(i)}$ einen Koffizienten > 1 haben, ferner hat das Element P^{ai} den Koeffizienten 1 in T^i. Andererseits folgt aus der mehrmaligen Anwendung des Haupthilfssatzes 1, daß T^i eine Linearkombination der Transitivitätsklassen von G ist. Nach vorigem tritt in dieser Linearkombination notwendig die Transitivitätsklasse $T(\mathrm{P}^{ai})$ mit dem Koeffizienten 1 auf, ferner kann diese nur Elemente von $T^{(i)}$ enthalten. Hieraus folgt wegen der Minimaleigenschaft von T, daß selbst $T^{(i)}$ eine Transitivitätsklasse von G sein muß, womit wir Hilfssatz 7 bewiesen haben.

Hilfssatz 8. *Enthält G eine endliche Transitivitätsklasse* ($\neq \mathrm{E}$), *so enthält es entweder ein Fixelement* ($\neq \mathrm{E}$) *oder eine selbstinverse Transitivitätsklasse mit nur zwei Elementen.*

Bezeichne nämlich T eine Transitivitätsklasse ($\neq \mathrm{E}$) von G mit minimaler Elementzahl t. Ist $t = 1$, so ist T ein Fixelement von G, und dann ist Hilfssatz 8 richtig. Im anderen Falle zeigen wir zunächst $t = 2$. Denn nehmen wir $t \geqq 3$ an. Da nach dem Haupthilfssatz 2 mit T zusammen auch $T^{(-1)}$ eine Transitivitätsklasse ist und dabei die Elementzahl unverändert bleibt, so darf (wegen $\mathrm{E} \bar{\in} T$)

$$\mathrm{P}^a,\ \mathrm{P}^b \in T \quad (a \neq b;\ \ a, b > 0) \tag{22}$$

angenommen werden. Aus Hilfssatz 6 folgt, daß $T^{(a)}$, $T^{(b)}$ auch Transitivitätsklassen von G sind. Beide enthalten P^{ab}, und so sind sie gleich. Dies ist wegen $a \neq b$ ein offenbarer Widerspruch, womit wir $t = 2$ bewiesen haben. Hiernach gilt (mit anderen a, b)

$$T = \mathrm{P}^a + \mathrm{P}^b \qquad (a \neq b;\ a,\ b \neq 0)\,. \tag{23}$$

Wegen $T^2 = \mathrm{P}^{2a} + \mathrm{P}^{2b} + 2\,\mathrm{P}^{a+b}$ folgt aus Haupthilfssatz 1, daß P^{a+b} (eine Transitivitätsklasse d. h.) ein Fixelement sein muß. Da aber jetzt (wegen $t = 2$) kein Fixelement ($\neq \mathrm{E}$) in G vorhanden ist, so ist $\mathrm{P}^{a+b} = \mathrm{E}$, $b = -a$, $T = P^a + P^{-a}$ ($a \neq 0$). Dies ist eben die allgemeine Form einer selbstinversen Transitivitätsklasse mit zwei Elementen, womit wir Hilfssatz 8 bewiesen haben.

Hilfssatz 9. *Ist T eine selbstinverse Transitivitätsklasse in G mit zwei Elementen, so sind es auch alle $T^{(i)}$* ($i = 1,2, \ldots$).

Denn wir dürfen

$$T = \mathrm{P}^a + \mathrm{P}^{-a} \quad (a > 0) \tag{24}$$

setzen. In T^i ($i > 1$) kommen P^{ai}, P^{-ai} mit der Multiplizität 1 vor, die übrigen Elemente haben eine Multiplizität > 1, und so folgt aus Haupthilfssatz 1, daß entweder $\mathrm{P}^{ai} + \mathrm{P}^{-ai} = T^{(i)}$ eine Transitivitätsklasse ist oder P^{ai}, P^{-ai} Fixelemente sind. Wir haben zu zeigen, daß stets der erste Fall zutrifft. Im anderen Falle gäbe es ein i (> 1) so, daß

$$\mathrm{P}^a + \mathrm{P}^{-a},\ \mathrm{P}^{a(i-1)} + \mathrm{P}^{-a(i-1)},\ \mathrm{P}^{ai},\ \mathrm{P}^{-a}$$

lauter Transitivitätsklassen sind. Dann gilt nach (6_1)

$$\mathrm{P}^{ai} = (\mathrm{P}^{ai})^{\Pi} = (\mathrm{P}^{a})^{\Pi}\,(\mathrm{P}^{a(i-1)})^{\Pi^{\mathrm{P}^a}} = \mathrm{P}^{-a}\,\mathrm{P}^{\pm a(i-1)}.$$

Mit diesem Widerspruch haben wir Hilfssatz 9 bewiesen.

Hilfssatz 10. *Kommt in G ein Fixelement ($\neq \mathrm{E}$) vor, so enthält G keine selbstinversen Transitivitätsklassen mit zwei Elementen.*

Denn nehmen wir an (vgl. Hilfssatz 1), daß $\{\mathrm{P}^r\}$ $(r>0)$ die Gruppe der Fixelemente von G ist und dabei G auch noch eine Transitivitätsklasse T mit zwei Elementen enthält. Diese nehmen wir wieder in der Form (24) an. Nach Hilfssatz 9 ist auch $T^{(r)} = \mathrm{P}^{ar} + \mathrm{P}^{-ar}$ eine Transitivitätsklasse, obwohl doch P^{ar}, P^{-ar} Fixelemente sind. Dieser Widerspruch beweist Hilfssatz 10.

Hilfssatz 11. *Gibt es in G eine selbstinverse Transitivitätsklasse mit zwei Elementen, so lassen sich alle selbstinversen Transitivitätsklassen mit zwei Elementen in der Form $T^{(i)}$ $(i = 1, 2, \ldots)$ angeben.*

Nehmen wir nämlich an, daß

$$T = \mathrm{P}^a + \mathrm{P}^{-a} \quad (a>0), \quad T_1 = \mathrm{P}^b + \mathrm{P}^{-b} \quad (b>0)$$

zwei verschiedene Transitivitätsklassen von G sind und dabei a minimal ist. Wegen Hilfssatz 9 genügt es $a|b$ zu beweisen. Es gilt eine Gleichung

$$b = ak + r \qquad (0 \leqq r < a; \quad k \geqq 1)$$

mit ganzen k, r. Nach (20) und Hilfssatz 9 gilt dann für $\varepsilon = \pm 1$ mit geeigneten $\varrho = \pm 1$, $\sigma = \pm 1$

$$(\mathrm{P}^{\varepsilon r})^{\Pi} = (\mathrm{P}^{\varepsilon b - \varepsilon a k})^{\Pi} = \mathrm{P}^{\varrho b}\,\mathrm{P}^{\varepsilon a k} = \mathrm{P}^{-\varepsilon b}\,\mathrm{P}^{\sigma a k}.$$

Hieraus folgt $(\varrho + \varepsilon) b = (\sigma - \varepsilon) a k$. Dies ergibt $\varrho = -\varepsilon$, $\sigma = \varepsilon$, also nach obigem

$$(\mathrm{P}^{\varepsilon r})^{\Pi} = \mathrm{P}^{-\varepsilon b + \varepsilon a k} = \mathrm{P}^{-\varepsilon r}.$$

Dies bedeutet, daß $\mathrm{P}^r + \mathrm{P}^{-r}$ eine Transitivitätsklasse von G ist, und so muß wegen der Minimaleigenschaft von a gewiß $r = 0$, $a|b$ sein. Wir haben Hilfssatz 11 bewiesen.

Hilfssatz 12. *Eine endliche Transitivitätsklasse von G kann nur ein Fixelement oder selbstinvers mit zwei Elementen sein.*

Denn nehmen wir an, daß es in G eine endliche Transitivitätsklasse

$$T = \mathrm{P}^{a_1} + \ldots + \mathrm{P}^{a_k} \quad (k \geqq 2) \tag{25}$$

gibt, die weder ein Fixelement noch selbstinvers mit zwei Elementen ist. Schon aus der schwächeren Annahme $T \neq \mathrm{E}$ folgt nach den Hilfssätzen 8, 10,

11 die Existenz von zwei eindeutig bestimmten ganzen Zahlen $d\,(>0)$, $\varepsilon\,(=\pm 1)$ mit

$$(26)\qquad (P^{di})^{\Pi} = P^{\varepsilon d i} \qquad (i = 0, \pm 1, \pm 2, \dots),$$

$$(27)\qquad (P^{i} + P^{-i})^{\Pi} \neq P^{i} + P^{-i} \qquad (d \nmid i)$$

(Im Fall $\varepsilon = 1$ ist nämlich $\{P^d\}$ die Gruppe der Fixelemente von G, im Fall $\varepsilon = -1$ sind die $P^{di} + P^{-di}$ $(i > 0)$ die sämtlichen selbstinversen Transitivitätsklassen mit zwei Elementen.) Bemerke man noch, daß nach diesen Definitionen gewiß

$$(28)\qquad d \nmid a_1, \dots, a_k$$

gilt.

Da nach (26), (6_1) und (5_1)

$$(P^{r+di})^{\Pi} = (P^{r})^{\Pi}\,(P^{\pm di})$$

gilt und der Exponent $r + di$ für $r\,(= 0, \dots, d-1)$, $i\,(= 0, \pm 1, \pm 2, \dots)$ alle ganzen Zahlen durchläuft, so folgt hieraus die Existenz einer Konstanten $\varkappa\,(>0)$ mit

$$(29)\qquad (P^{i})^{\Pi} = P^{\pm i + i^*} \qquad (i = 0 \pm 1, \pm 2, \dots;\ |i^*| < \varkappa).$$

Bezeichne p eine Primzahl. Aus der bekannten Teilbarkeitseigenschaft der Polynomialkoeffizienten folgt, daß in T^p die Elemente von $T^{(p)}$ eine Multiplizität $\equiv 1 \pmod{p}$, die übrigen Elemente eine durch p teilbare Multiplizität haben. Dies ergibt nach dem Haupthilfssatz 1 sofort

$$(T^{(p)})^{\Pi} = T^{(p)},$$

d. h. nach (25)

$$(30)\qquad (P^{pa_j})^{\Pi} = P^{pa_{j'}} \qquad (j = 1, \dots, k),$$

wobei j' eine (von p abhängige) Permutation der j bezeichnet. Aus (29), (30) folgt aber, daß für genügend große p (nämlich gewiß für $p > \varkappa$) $a_{j'} = \pm a_j$ sein muß. Wendet man dies auch auf $T^{(-1)}$ statt T an (vgl. Haupthilfssatz 2), so hat man das Resultat:

$$(31)\qquad (P^{\pm pa_j})^{\Pi} = P^{\pm pa_j} \qquad (j = 1, \dots, k;\ p > \varkappa),$$

wobei das Vorzeichen links beliebig wählbar, rechts vorläufig nicht näher bestimmt ist.

Da aber nach Hilfssatz 1 mit einem Element von G zusammen auch sein Inverses ein Fixelement ist, so müssen in (31) für jedes feste Paar p, j stets entweder beiderseits gleiche oder beiderseits ungleiche Vorzeichen stehen. Das bedeutet, daß statt (31) genauer

$$(P^{pa_j} + P^{-pa_j})^{\Pi} = P^{pa_j} + P^{-pa_j}$$

gilt. Wegen (28) ist das für $(p, d) = 1$ ein Widerspruch zu (27), womit wir Hilfssatz 12 bewiesen haben.

Hilfssatz 13. *Enthält Γ ein Fixelement ($\neq \mathsf{E}$), so enthält G nur endliche Transitivitätsklassen.*

Bezeichne nämlich C ($\neq \mathsf{E}$) ein Fixelement von Γ. Nach Hilfssatz 5 bewirkt C einen Automorphismus von G. Das bedeutet $\mathrm{A}^{\mathsf{C}} = \mathrm{A}^{\varepsilon}$ ($\mathrm{A} \in G$) mit konstantem $\varepsilon = \pm 1$, also gilt $\mathrm{A}^{\mathsf{C}^2} = \mathrm{A}$, d. h. C^2 ist ein identischer Operator. Gewiß ist dann die Faktorgruppe Γ/Γ^0 in (16) endlich, woraus nach der dort gemachten Bemerkung die Endlichkeit aller Transitivitätsklassen von G folgt. Das beweist Hilfssatz 13.

Hilfssatz 14. *Gibt es in Γ keine endlichen Transitivitätsklassen ($\neq \mathsf{E}$), so folgt aus jeder Gleichung $\mathsf{A}^{\mathrm{A}} = \mathsf{A}^{-1}$ ($\mathrm{A} \in G$, $\mathsf{A} \in \Gamma$, $\mathsf{A} \neq \mathsf{E}$) auch das Bestehen von $\mathrm{A}^{\mathsf{A}} = \mathrm{A}^{-1}$.*

Nach der Annahme gilt nämlich $\mathsf{A}^{\mathrm{A}}\mathsf{A} = \mathsf{E}$. Wendet man beiderseits den Operator A an, so folgt nach Formel (6_2)

$$(\mathsf{A}^{\mathrm{A}})^{\mathrm{A}^{\mathsf{A}}}\, \mathsf{A}^{\mathrm{A}} = \mathsf{E},$$

also nach (5_2) und voriger Gleichung:

$$\mathsf{A}^{\mathrm{A}^{\mathsf{A}}\mathrm{A}} = \mathsf{A}.$$

Ist die Behauptung falsch, d. h. $\mathrm{A}^{\mathsf{A}}\mathrm{A} \neq \mathrm{E}$, so besagt die erhaltene Gleichung, daß die Transitivitätsklasse $\mathsf{T}(\mathsf{A})$ von Γ endlich ist. Dieser Widerspruch beweist Hilfssatz 14.

§ 4. Der Fall zyklischer Gruppen G, Γ mit $O(G) = \infty$, $O(\Gamma) = \mu$

Wir betrachten weiter auch die zyklischen Gruppen $G = \{\mathrm{P}\}$, $\Gamma = \{\Pi\}$, nehmen aber jetzt an, daß G unendlich, Γ ($\neq \mathsf{E}$) endlich ist und setzen

(32) $$O(\Gamma) = \mu (> 1).$$

Anschließend bestimmen wir in diesem Spezialfall alle Funktionenpaare (2) aus dem Satz von Zappa.

Wegen (32) ist die Permutationsgruppe $[\Gamma]$ endlich, und so müssen alle Transitivitätsklassen von G endlich sein. Wegen der Hilfssätze 12, 10 gilt also mit einem festen ε

(33) $$(\mathrm{P}^i)^{\Pi} = \mathrm{P}^{\varepsilon i} \qquad (\varepsilon = \pm 1;\ i = 0, \pm 1, \pm 2, \ldots).$$

Erster Fall: $\varepsilon = 1$. Wir werden leicht zeigen, daß jezt die gesuchten Funktionenpaare (2) durch[6]

$$(P^i)^{\Pi^k} = P^i\,,\quad (\Pi^k)^{P^i} = \Pi^{k\,g^i} \tag{34}\qquad ((g,\mu)=1)$$

mit einer sonst beliebigen ganzen Zahl g gegeben sind.

Aus (33) und $\varepsilon = 1$ folgt nämlich (34_1). Wir setzen $\Pi^P = \Pi^g$. Nach Hilfssatz 5 (angewendet für G, Γ statt Γ, G) gilt dann (34_2) zunächst für $i = 1$, also allgemeiner für $i \geqq 0$. Man sieht auch, daß $(g, \mu) = 1$ sein muß, denn sonst wäre [P] keine Permutation von Γ. Endlich gilt (34_2) wegen (5_2) offenbar auch für $i < 0$.

Umgekehrt sieht man leicht, daß (34) stets eine Lösung von (4), (5), (6) ist, womit wir die Behauptung bewiesen haben.

Zweiter Fall: $\varepsilon = -1$. Aus (33) folgt sofort

$$(P^i)^{\Pi^k} = P^{i(-1)^k}. \tag{35}$$

Dies ergibt wegen $\Pi^\mu = E$ für $k = \mu$ mit Notwendigkeit

$$2 | \mu\,. \tag{36}$$

Um auch die Funktion (2_2) zu bestimmen, gehen wir so vor. Zunächst folgt aus (35) $P^{\Pi^2} = P$, also wegen (6_2)

$$(\Pi^{k+2})^P = (\Pi^k)^P\,(\Pi^2)^P\,.$$

Mit Induktion schließt man hieraus auf

$$(\Pi^{k+2r})^P = (\Pi^k)^P\,((\Pi^2)^P)^r\,, \tag{37}$$

gültig zunächst für $r \geqq 0$, offenbar aber auch für $r < 0$. Insbesondere für $k = 0$, $r = \frac{\mu}{2}$ bekommen wir $((\Pi^2)^P)^{\frac{\mu}{2}} = E$ und so gilt

$$(\Pi^2)^P = \Pi^{2a} \tag{38}$$

mit einer ganzen Zahl a. Wir setzen

$$\Pi^P = \Pi^{a+b} \tag{39}$$

mit einer passenden ganzen Zahl b.

Bezeichnen wir für ganze x mit $\bar{x}$ den kleinsten nichtnegativen Rest von x mod 2, d. h. wir setzen

$$\bar{x} = \begin{cases} 0 & (2 \mid x), \\ 1 & (2 \nmid x). \end{cases} \tag{40}$$

Mit dieser Bezeichnung kommen wir von (37), (38), (39) leicht zur Formel

$$(\Pi^k)^P = \Pi^{ak + b\bar{k}}\,. \tag{41}$$

[6] Selbstverständlich kommt g^i nur mod μ in Betracht.

Der Exponent auf der rechten Seite muß wegen (36) für passende k auch ungerade Werte annehmen, und so folgt hieraus wegen

$$a k + b \bar{k} \equiv (a + b) k \pmod 2$$

mit Notwendigkeit

$$2 \nmid a + b . \tag{42}$$

Die a, b sind nicht unabhängig, wie wir das hier prüfen werden. Aus diesem Zweck berechnen wir nach (6_2) und (35):

$$(\pi^{k+1})^{\mathrm{P}} = (\pi^k)^{\mathrm{P}^{\pi}} \pi^{\mathrm{P}} = (\pi^k)^{\mathrm{P}-1} \pi^{\mathrm{P}} .$$

Wegen (41) und $\overline{k + 1} = 1 - \bar{k}$ schreibt sich dies als

$$\pi^{a(k+1) + b(1-\bar{k})} = (\pi^k)^{\mathrm{P}-1} \pi^{a+b},$$

d. h.

$$\pi^{a k - b \bar{k}} = (\pi^k)^{\mathrm{P}-1} .$$

Der Exponent auf der linken Seite ist wegen (42) kongruent $k \pmod 2$. Wendet man also beiderseits den Operator P an, so folgt aus (41), (5_2)

$$\pi^{a(a k - b\bar{k}) + b\bar{k}} = \pi^k .$$

Dies ergibt

$$a(a k - b \bar{k}) + b \bar{k} \equiv k \pmod \mu .$$

Insbesondere für $k = 2, 1$ folgt hieraus

$$a^2 - 1 \equiv 0 \left(\mathrm{mod}\ \frac{\mu}{2}\right), \quad a^2 - 1 \equiv (a - 1) b \pmod \mu . \tag{43}$$

Jetzt sind wir schon in der Lage

$$(\pi^k)^{\mathrm{P}^i} = \pi^{k + (a-1)\bar{i} k + b i \bar{k}} \tag{44}$$

zu beweisen. Dies lautet für $i + 1$ statt i so:

$$(\pi^k)^{\mathrm{P}^{i+1}} = \pi^{k + (a-1)(1-\bar{i}) k + b(i+1)\bar{k}} . \tag{45}$$

Da (44) für $i = 0$ richtig ist, werden wir (44) für $i = 0, \pm 1, \pm 2, \ldots$ bewiesen haben, wenn wir zeigen, daß (44), (45) auseinander folgen. Es genügt zu zeigen, daß die rechte Seite von (44) nach Ausübung des Operators P in die von (45) übergeht. Da nach (42) der Exponent rechts in (44) kongruent k mod 2 ist, so kommen wir nach (41) zu

$$\pi^{a(k + (a-1)\bar{i} k + b i \bar{k}) + b \bar{k}} .$$

Wir haben zu zeigen, daß die beiden Exponenten hier und rechts in (45) mod μ kongruent sind, d. h. für ihre Differenz

$$(a^2 - 1) \bar{i} k + (a - 1) b i \bar{k} \equiv 0 \pmod \mu$$

gilt. Diese Kongruenz ist wegen (43) in der Tat befriedigt, womit wir (44) bewiesen haben.

Leicht überzeugt man sich, daß umgekehrt das Funktionenpaar (35), (44) den Gleichungen (4), (5), (6) genügt, wenn (36), (42), (43) erfüllt sind. Mit dem vorigen Resultat zusammen haben wir den folgenden:

Satz 1. *Es seien G, Γ zyklische Gruppen mit den Erzeugenden* P, Π *von unendlicher bzw. der endlichen Ordnung μ. Die sämtlichen Lösungen (2) der Bedingungsgleichungen (4), (5), (6) von* Zappa *sind*

$$(46)\qquad (\mathrm{P}^i)^{\Pi^k} = \mathrm{P}^i, \qquad (\Pi^k)^{\mathrm{P}^i} = \Pi^{kg^i} \qquad ((g, \mu) = 1)$$

und im Falle $2 \mid \mu$ noch

$$(47)\qquad (\mathrm{P}^i)^{\Pi^k} = \mathrm{P}^{i(-1)^k}, \quad (\Pi^k)^{\mathrm{P}^i} = \Pi^{k+(a-1)\bar{i}k+bi\bar{k}}$$

$$\left|2 + a + b,\ \frac{\mu}{2} \,\middle|\, a^2 - 1,\ \mu \,\middle|\, a^2 - 1 + (a^2 - 1)b\right|.$$

Bemerkung. Nach dem Satz von Zappa liefern (46), (47) alle Gruppen $\mathfrak{G} = G\,\Gamma$ in diesem Falle. Zur vollständigen Beantwortung dieser Frage ist natürlich nötig, daß man die nichtisomorphen unter ihnen bestimmt. Das tun wir im § 6, wo wir auch die näheren Eigenschaften dieser Gruppen untersuchen werden. Hier wollen wir aber auf den (in der Einleitung schon vorhergesagten) Umstand aufmerksam machen, daß insbesondere die Lösung (47) von (4), (5), (6) durch die beiden „Anfangswerte" $\mathrm{P}^{\Pi} = \mathrm{P}^{-1}$, $\Pi^{\mathrm{P}} = \Pi^{a+b}$ noch nicht bestimmt ist, erst die Hinzunahme des weiteren „Anfangswertes" $(\Pi^2)^{\mathrm{P}} = \Pi^{2a}$ schafft Eindeutigkeit.

§ 5. Der Fall zyklischer Gruppen G, Γ mit $O(G) = O(\Gamma) = \infty$

Wir wollen jetzt den Fall von zwei unendlichen zyklischen Gruppen $G = \{\mathrm{P}\}$, $\Gamma = \{\Pi\}$ betrachten. In diesem Fall können wir nur diejenigen Lösungen (2) von (4), (5), (6) bestimmen, für die es unter den zugehörigen Transitivitätsklassen ($\neq \mathrm{E}$ bzw. E) von G und Γ mindestens eine endliche gibt. Für die etwaigen übrigen Lösungen werden wir uns mit einigen Bemerkungen begnügen müssen, uns scheint, daß solche Lösungen überhaupt nicht vorhanden sind.

Betrachten wir eine Lösung A^{A}, A^{A} von (4), (5), (6) unter der gesagten Einschränkung. Wir unterscheiden die folgenden Fälle 1), 2), 3), wovon sich 2) in die weiteren Fälle 21), 22) spalten wird:

Fall 1): Beide Gruppen G, Γ enthalten Fixelemente außer E bzw. E. Nach Hilfssatz 13 (angewendet auf G, Γ und auf Γ, G) enthalten G, Γ nur endliche Transitivitätsklassen, woraus nach den Hilfssätzen 12, 10 folgt,

daß G und ebenso auch Γ überhaupt nur Fixelemente enthält. Mit anderen Worten gilt jetzt $A^{\mathsf{A}} = A$, $\mathsf{A}^{A} = \mathsf{A}$ unbeschränkt, das ist die identische Lösung von (4), (5), (6).

Fall 2): G enthält keine Fixelemente ($\neq$ E), Γ enthält Fixelemente ($\neq \mathsf{E}$). Aus letzterem folgt nach Hilfssatz 13 wieder, daß G aus lauter endlichen Transitivitätsklassen besteht; diese müssen jetzt nach den Hilfssätzen 12, 10 selbstinvers sein mit zwei Elementen. Das bedeutet, daß unbeschränkt $A^{\Pi} = A^{-1}$ gilt. Dann ist Π^2 ein identischer Operator für G, und so folgt aus Hilfssatz 6 (angewendet für Γ, G statt G, Γ), daß P als Operator einen Automorphismus der Gruppe $\{\Pi^2\}$ bewirkt. Hiernach gilt $(\Pi^{2i})^{P} = \Pi^{2i\varepsilon}$ mit einem festen $\varepsilon = \pm 1$. Der Fall $\varepsilon = -1$ ist unmöglich, denn dann hätte Γ selbstinverse Transitivitätsklassen mit zwei Elementen, was jetzt wegen Hilfssatz 10 unmöglich ist. Also gilt $\varepsilon = 1$, und dies bedeutet, daß Π^2 ein Fixelement von Γ ist.

Fall 21): Π ist ein Fixelement von Γ. Mit obigem zusammen gilt dann

(48) $$(P^i)^{\Pi^k} = P^{i(-1)^k}, \quad (\Pi^k)^{P^i} = \Pi^k .$$

Umgekehrt sieht man, daß dies eine Lösung von (4), (5), (6) ist.

Fall 22): Π ist kein Fixelement von Γ. Da jetzt $\{\Pi^2\}$ die Gruppe der Fixelemente von Γ ist, so muß Π^P außerhalb $\{\Pi^2\}$ liegen. Deshalb können wir

$$\Pi^P = \Pi^{1+2g}$$

setzen mit einem ganzen g ($\neq 0$). Aus (6_2) folgt wegen $(\Pi^{2k})^P = \Pi^{2k}$:

$$(\Pi^{2k+1})^P = \Pi^{2k}\Pi^P = \Pi^{2k+1+2g}.$$

Beide lassen sich (mit Hilfe des in (40) eingeführten Zeichens $\bar{x}$) vereinigen als

$$(\Pi^k)^P = \Pi^{k+2g\bar{k}}.$$

Dies läßt sich wegen $k + 2g\bar{k} \equiv k \pmod 2$ leicht auf P^i (statt P) verallgemeinern, und so kommen wir zu

(49) $$(P^i)^{\Pi^k} = P^{i(-1)^k}, \quad (\Pi^k)^{P^i} = \Pi^{k+2gi\bar{k}} \qquad (g \neq 0).$$

Umgekehrt sieht man leicht, daß (49) stets eine Lösung von (4), (5), (6) ist.

Fall 3): Weder G noch Γ enthält Fixelemente ($\neq$ E bzw. E). Wegen der oben getroffenen Einschränkung dürfen wir aus Symmetriegründen annehmen, daß G endliche Transitivitätsklassen ($\neq$ E) enthält. Diese sind nach Hilfssatz 12 selbstinvers mit zwei Elementen, also gibt es in G ein Element A ($\neq$ E) mit

$$A^{\Pi^k} = A^{(-1)^k}.$$

Wir berechnen nach (20_2):

$$(\Pi^{2k+1})^{A} = \begin{cases} (\Pi^{2k})^{A^{\Pi}}\,\Pi^{A} = (\Pi^{2k})^{A^{-1}}\,\Pi^{A}, \\ (\Pi^{2k})^{A}\,\Pi^{A^{\Pi^{2k}}} = (\Pi^{2k})^{A}\,\Pi^{A}. \end{cases}$$

Hieraus folgt $(\Pi^{2k})^{A^{-1}} = (\Pi^{2k})^{A}$ also nach (5_2)

$$(\Pi^{2k})^{A^2} = \Pi^{2k}.$$

Dies hat zur Folge, daß Π^{2k} in eine endliche Transitivitätsklasse gehört. Insbesondere folgt dann aus der hiermit erwiesenen Existenz von endlichen Transitivitätsklassen ($\neq \mathsf{E}$) auch in Γ, daß das vorhergesagte auch für P statt Π gilt, und so gilt nach Hilfssatz 12

(50) $$(P^{2i})^{\Pi^k} = P^{2i(-1)^k}, \quad (\Pi^{2k})^{P^i} = \Pi^{2k(-1)^i}.$$

Wir zeigen, daß sogar

(51) $$(P^{i})^{\Pi^k} = P^{i(-1)^k}, \quad (\Pi^{k})^{P^i} = \Pi^{k(-1)^i}$$

gilt. Wegen (50) muß nämlich

$$P^{\Pi} = P^{1+2a}, \quad \Pi^{P} = \Pi^{1+2b}$$

gelten mit irgendwelchen ganzen a, b. Wir berechnen nach (5_1), (6_1) und (50):

$$P^{\Pi^2} = (P^{1+2a})^{\Pi} = P^{\Pi}\,(P^{2a})^{\Pi^{P}} = P^{\Pi}\,(P^{2a})^{\Pi^{1+2b}} = P^{1+2a}P^{-2a} = P.$$

Hiernach gehört P in eine endliche Transitivitätsklasse, und so folgt aus Hilfssatz 12, daß $P + P^{-1}$ eine Transitivitätsklasse sein muß. Nach Hilfssatz 9 gilt also (51_1) zunächst für $k = 1$, dann aber auch allgemein. Aus Symmetriegründen ist auch (51_2) richtig.

Umgekehrt sieht man sofort, daß (51) eine Lösung von (4), (5), (6) ist.

Da (48) der Fall $g = 0$ von (49) ist, so sammeln wir die Resultate von diesem Paragraphen im folgenden:

Satz 2. *Es seien G, Γ unendliche zyklische Gruppen mit den Erzeugenden P bzw. Π. Diejenigen Lösungen (2) der Bedingungsgleichungen (4), (5), (6) von* Zappa, *für die es ein Paar X ($\neq E$), X ($\neq \mathsf{E}$) mit $X^{\mathsf{X}} = X$ oder $\mathsf{X}^{X} = \mathsf{X}$ gibt,*[7] *sind (von der identischen Lösung $A^{\mathsf{A}} = A$, $\mathsf{A}^{A} = \mathsf{A}$ abgesehen) im wesentlichen nur die folgenden:*[8]

(52) $$(P^{i})^{\Pi^k} = P^{i(-1)^k}, \qquad (\Pi^{k})^{P^i} = \Pi^{k(-1)^i},$$

(53) $$(P^{i})^{\Pi^k} = P^{i(-1)^k}, \qquad (\Pi^{k})^{P^i} = \Pi^{k+2gi\bar{k}}.$$

[7] Es muß bemerkt werden, daß bei festem X die Existenz einer Lösung X ($\neq \mathsf{E}$) von $X^{\mathsf{X}} = X$ mit der Endlichkeit der Transitivitätsklasse $T(X)$ gleichbedeutend ist, und so wurden im Satz 2 eben diejenigen Lösungen (2) außer Acht gelassen, für die sowohl G als auch Γ lauter unendliche Transitivitätsklassen ($\neq E$ bzw. E) enthalten.

[8] Die Funktion $\bar{k}$ haben wir in (40) erklärt.

(Als nicht wesentlich neue Lösungen treten außerdem diejenigen hinzu, die aus (53) nach Vertauschung von P *mit* Π *entstehen.)*

Bemerkungen. Wir sehen, daß von den beiden Funktionen (53) trotz der Isomorphie $G \approx \Gamma$ die zweite verwickelter ist als die erste, auf diese Asymmetrie haben wir in der Einleitung schon hingespielt. Die näheren Eigenschaften der aus (52), (53) entspringenden Gruppen werden wir im § 6 untersuchen.

§ 6. Nähere Beschreibung der in den Sätzen 1, 2 gewonnenen Gruppen

Auf Grund des Satzes von ZAPPA liefern die vier Funktionenpaare (46), (47), (52), (53) je eine Gruppe, die wir am bequemsten mit $\mathfrak{G}_{46}$, $\mathfrak{G}_{47}$, $\mathfrak{G}_{52}$, $\mathfrak{G}_{53}$ bezeichnen. Und zwar sind diese Gruppen lauter ZAPPA-SZÉPschen Produkte $(G\Gamma =)\, \}\mathrm{P}\{\,\}\Pi\{$, wobei stets $O(\mathrm{P}) = \infty$, ferner in den ersten zwei Fällen $O(\Pi) = \mu$, in den letzten zwei Fällen $O(\Pi) = \infty$ ist. Wir wollen einiges über diese Gruppen kurz bemerken, führen aber die Rechnungen nur teils aus. Selbstverständlich kommen die Parameter g, a, b in $\mathfrak{G}_{46}$, $\mathfrak{G}_{47}$ nur mod μ in Betracht. Den Fall $\mu \mid g-1$ von $\mathfrak{G}_{46}$ schließen wir aus, da es sich dann um das direkte Produkt von G, Γ handelt. Ebenfalls schließen wir den Fall

$$\mu \mid 2b, \ a-1+b$$

von $\mathfrak{G}_{47}$ aus, da dann diese Gruppe sichtbar isomorph zum Fall $\mu \mid g-1$ von $\mathfrak{G}_{46}$ ist.

Wir werden sehen, daß die Abhängigkeit von $\mathfrak{G}_{47}$ von den beiden Parametern a, b nur scheinbar ist, statt ihrer werden wir mit einem (invarianten) Parameter d $\left(d \,\middle|\, \frac{\mu}{2}\right)$ auskommen. Auch der Parameter g in $\mathfrak{G}_{53}$ wird sich herauswerfen lassen, und zwar werden wir sehen, daß die Fälle $g = 0, 1$ schon alle nichtisomorphen $\mathfrak{G}_{53}$ erschöpfen.

Wir schicken voraus, daß nach (7) für eine beliebige Gruppe $\mathfrak{G} = G\Gamma$ im Satz von ZAPPA die zwei beliebigen Elemente $\mathrm{A}\,\mathsf{A}$, $\mathrm{B}\,\mathsf{B}$ dann und nur dann vertauschbar sind, wenn

$$\mathrm{A}\,\mathrm{B}^{\mathsf{A}} = \mathrm{B}\,\mathrm{A}^{\mathsf{B}}, \quad \mathsf{A}^{\mathrm{B}}\mathsf{B} = \mathsf{B}^{\mathrm{A}}\mathsf{A}$$

gelten. Ferner schreiben wir den Spezialfall von (7) hin:

$$(\mathrm{A}\,\mathsf{A})^2 = \mathrm{A}\,\mathrm{A}^{\mathsf{A}}\,\mathsf{A}^{\mathrm{A}}\mathsf{A}.$$

Die Kommutatorgruppe und das Zentrum von $\mathfrak{G}$ bezeichnen wir mit $\mathfrak{G}'$ bzw. $\mathfrak{G}^{\circ}$. Insbesondere für den uns interessierenden Fall $G = \}\mathrm{P}\{$, $\Gamma = \}\Pi\{$ setzen wir

$$(\mathrm{P}, \Pi) = \mathrm{P}^{-1}\,\Pi\,\mathrm{P}\,\Pi^{-1}.$$

Offenbar wird dann $\mathfrak{G}'$ durch die Konjugierten von (P, Π) erzeugt.

Nunmehr betrachten wir die Gruppen $\mathfrak{G}_{46}$, $\mathfrak{G}_{47}$, $\mathfrak{G}_{52}$, $\mathfrak{G}_{53}$ unter 1), 2), 3), 4):

1) Für $\mathfrak{G}_{46}$ ist Γ offenbar eine normale Untergruppe, diese umfaßt zugleich alle Elemente endlicher Ordnung. Man sieht leicht, daß zwei Parameterwerte g, g' dann und nur dann zu isomorphen Gruppen $\mathfrak{G}_{46}$ führen, wenn $g\,g' \equiv 1 \pmod{\mu}$ gilt. Es gelten

$$\mathfrak{G}'_{46} = \{\Pi^{g-1}\}\,,\quad \mathfrak{G}^{\circ}_{46} = \{\mathrm{P}^{O(g \,(\mathrm{mod}\, \mu))}\,,\ \Pi^{\mu(\mu, g-1)^{-1}}\}\,,$$

wobei $O(g \,(\mathrm{mod}\, \mu))$ die Ordnung von g mod μ bezeichnet.

2) Für $\mathfrak{G}_{47}$ sieht man vor allem, daß (47) gegenüber der Ersetzung von a, b durch $a + \frac{\mu}{2}$, $b + \frac{\mu}{2}$ invariant bleibt. Deshalb darf stets $2 \nmid a$ (also $2 \mid b$) angenommen werden. Weder G noch Γ ist normal in $\mathfrak{G}_{47}$, denn es gelten nach (47) (und (7))

$$\Pi\,\mathrm{P}\,\Pi^{-1} = \mathrm{P}^{-1}\,\Pi^{a-1+b}\,,\quad \mathrm{P}^{-1}\,\Pi = \mathrm{P}^{-2}\,\Pi^{a+b}\,,$$

nie kann aber $\Pi^{a-1+b} = \mathrm{E}$ sein, da hieraus $\mu \mid a - 1 + b$ (ferner wegen (47)) $\mu \mid 2b$ folgte, und diesen Fall haben wir oben ausgeschlossen. Leicht zeigt man

$$\mathfrak{G}'_{47} = \{\mathrm{P}^{-2}\,\Pi^{a-1+b}\,,\ \Pi^{2(a-\ )}\}\,,\quad \mathfrak{G}^{\circ}_{47} = \left\{\Pi^{\mu/\left(\frac{\mu}{2}, a-1\right)}\right\},$$

Aus letzterem folgt $O(\mathfrak{G}^{\circ}_{47}) = \left(\frac{\mu}{2}, a-1\right)$, und so ist

$$d = \left(\frac{\mu}{2}, a-1\right) \tag{54}$$

eine Invariante von $\mathfrak{G}_{47}$. Dabei gilt (wegen $2 \nmid a$):

$$d \,\Big|\, \frac{\mu}{2}\,,\qquad \left(\frac{\mu}{2}, 2\right) \Big|\, d\,. \tag{55}$$

Wir bezeichnen mit d' den komplementären Teiler:

$$d\,d' = \frac{\mu}{2}\,. \tag{56}$$

Nunmehr zeigen wir, daß man alle nichtisomorphen $\mathfrak{G}_{47}$ schon bei der Einschränkung

$$b = d',\ 2d' \text{ oder } 4d' \tag{57}$$

erfaßt. Denn setzen wir

$$\mathrm{P}' = \mathrm{P}\,\Pi^{2x}\,,\quad \Pi' = \mathrm{P}^{2y}\,\Pi^{z-by} \qquad ((z, \mu) = 1)\,.$$

Man sieht leicht nach, daß P', Π' ein mit P, Π gleichberechtigtes Elementenpaar ist (darunter verstehen wir, daß $\mathfrak{G}_{47} = \{\mathrm{P}'\}\,\{\Pi'\}$ mit $O(\mathrm{P}') = \infty$, $O(\Pi) = \mu$

gilt). Ferner berechnet man

$$\Pi'^2 \mathrm{P}' = \mathrm{P}' \Pi'^{2a}, \quad \Pi' \mathrm{P}' = \mathrm{P}'^{-1} \Pi'^{a+b'}$$

$$(zb' \equiv (a+1)(2x + (a-1)y) + b \pmod{\mu}).$$

Andererseits gilt

$$\Pi^2 \mathrm{P} = \mathrm{P} \Pi^{2a}, \quad \Pi \mathrm{P} = \mathrm{P}^{-1} \Pi^{a+b}.$$

Vergleicht man diese miteinander, so sieht man, daß bei Ersetzung von P, Π durch P', Π' das Parameterpaar a, b in a, b' übergeht. Ferner läßt sich für passende x, y, z erreichen, daß

$$b' = (2(a+1), b, \mu).$$

Andererseits folgt aus (54)

$$db' = \left(a-1, \frac{\mu}{2}\right)(2(a+1), b, \mu) = \left(2(a^2-1), b(a-1), \mu(a-1), 2\mu, \frac{b}{2}\mu, \frac{\mu}{2}\mu\right) =$$

$$= \frac{\mu}{2}, \mu, \text{ oder } 2\mu.$$

Hieraus und aus (56) folgt, daß für b' nur die drei Werte d', $2d'$, $4d'$ möglich sind, womit die Behauptung über (57) bewiesen ist. Wegen $\frac{\mu}{2} \mid a^2 - 1$ und (54), (56) muß auch

(58) $$d \mid a-1, \qquad d' \mid a+1$$

gelten. Dies und (54) ergeben auch

(59) $$(d, d') = 1 \text{ oder } 2.$$

Nach diesen haben wir folgendes gewonnen. Sämtliche $\mathfrak{G}_{47}$ (bis auf Isomorphie) gewinnt man auch schon so, daß man $\frac{\mu}{2}$ nach (56), (59) zerlegt, a und b nach (58) und (57) wählt, ferner auch auf $(\mu =)\, 2dd' \mid (a-1)(a+1+b)$ achtet. Es wäre leicht unter den so verbliebenen $\mathfrak{G}_{47}$ die nichtisomorphen herauszuscheiden, wovon wir aber Abstand nehmen. Wir wollten bloß zeigen, daß ein nur von einem Parameter $d\left(d \,\middle|\, \frac{\mu}{2}\right)$ abhängiges volles System von nichtisomorphen $\mathfrak{G}_{47}$ sich angeben läßt. Wir bemerken noch, daß $\{\mathrm{P}^2, \Pi^2\}$ eine Abelsche normale Untergruppe von $\mathfrak{G}_{47}$ vom Index 4 ist. Die Untergruppe $\{\mathrm{P}, \Pi^2\}$ ist vom Index 2 (also auch normal), ihre Elemente $(\neq \mathrm{E})$ sind von unendlicher Ordnung, dagegen besteht die Nebengruppe $\{\mathrm{P}, \Pi^2\}\,\Pi$ aus lauter Elementen endlicher Ordnung.

3) Für $\mathfrak{G}_{52}$ gilt

$$\mathfrak{G}'_{52} = \{\mathrm{P}^2 \Pi^2, \mathrm{P}^4\}, \quad \mathfrak{G}''_{52} = \mathrm{E}.$$

Keine der Untergruppen G, Γ ist normal, die Untergruppe $\{\mathrm{P}^2, \Pi^2\}$ ist wieder Abelsch und normal vom Index 4, die Elemente der Nebengruppe $\{\mathrm{P}^2, \Pi^2\}\,\mathrm{P}\Pi$

sind von 2-ter Ordnung, die übrigen Elemente ($\neq \mathrm{E}$) sind von unendlicher Ordnung.

4) Für $\mathfrak{G}_{53}$ gilt

$$\mathfrak{G}'_{53} = \{\mathrm{P}^{-2}\,\Pi^{2g}\}\,, \quad \mathfrak{G}^{\circ}_{53} = \{\Pi^2\}\,.$$

Wieder ist $\{\mathrm{P}^2, \Pi^2\}$ eine Abelsche normale Untergruppe. Wir zeigen, daß die beiden Fälle $g = 0, 1$ schon alle nichtisomorphen $\mathfrak{G}_{53}$ sind. Setzt man nämlich $\mathrm{P}' = \mathrm{P}\,\Pi^{2x}$, so ist P' ein mit P gleichberechtigtes Element (d. h. es gilt $\mathfrak{G}_{53} = \{\mathrm{P}'\}\,\{\Pi\}$ mit $O(\mathrm{P}') = \infty$). Wir berechnen:

$$\Pi\,\mathrm{P}' = \mathrm{P}'^{-1}\,\Pi^{1+2(g+2x)}\,.$$

Andererseits gilt

$$\Pi\,\mathrm{P} = \mathrm{P}^{-1}\,\Pi^{1+2g}\,.$$

Diese zeigen schon, daß sich der Parameter g bei passender Wahl von x auf einen der Werte 0, 1 reduzieren läßt. Im Fall $g = 0$ ist unter den Untergruppen G, Γ nur die zweite normal, im Fall $g = 1$ ist keine dieser Gruppen normal. Über den Fall $g = 1$ bemerken wir, daß dann $\mathfrak{G}_{53}$ auch Elemente zweiter Ordnung enthält, diese sind nämlich die und nur die $\mathrm{P}^i\,\Pi^{-i}$ ($i = \pm 1, \pm 3, \pm 5, \ldots$), die übrigen Elemente sind (außer dem Einselement) von endlicher Ordnung. Wir wollen auf folgenden interessanten Umstand aufmerksam machen. Betrachten wir wieder alle Parameterwerte g ($= 0, \pm 1, \ldots$), fassen aber z. B. nur die ungeraden g ins Auge. (Nach obigem sind die entsprechenden G_{53} alle miteinander isomorph.) Die Transitivitätsklassen von Γ sind die folgenden: Zunächst sind die Π^{2} ($k = 0, \pm 1, \ldots$) lauter Fixelemente, hierzu kommen noch $|g|$ (unendliche) Transitivitätsklassen von der Form Π^{a+2gi} ($i = 0, \pm 1, \ldots$), wobei a eine der Zahlen $1, 3, \ldots, 2|g| - 1$ ist. Dieses Beispiel zeigt uns, daß in einem ZAPPA—SZÉPschen Produkt $\mathfrak{G} = G\,\Gamma$ die Natur der Transitivitätsklassen der Faktoren G, Γ sich stark ändern kann, wenn man auf eine äquivalente Repräsentation übergeht. (Auch im Fall 2) sehen die Transitivitätsklassen von Γ ziemlich bunt aus, wie man das von (47_2) abliest.)

§ 7. Bemerkungen über den Ausnahmefall von § 5

Wieder betrachten wir den Fall von zwei unendlichen zyklischen Gruppen $G = \{\mathrm{P}\}$, $\Gamma = \{\Pi\}$. Im Satz 2 haben wir bei der Bestimmung der Lösungen A^{A}, A^{A} der Bedingungsgleichungen (4), (5), (6) von ZAPPA den Fall mit

$$\mathrm{A}^{\mathsf{A}} \neq \mathrm{A}\,, \quad \mathsf{A}^{\mathrm{A}} \neq \mathsf{A} \qquad (\mathrm{A} \neq \mathrm{E}\,,\ \mathsf{A} \neq \mathsf{E}) \tag{60}$$

ausgenommen. Wir wollen einiges über diesen Ausnahmefall bemerken, um hiermit der Lösung dieses restlichen Problems näher zu kommen.

Wir wiederholen, daß nach[7] die fraglichen Lösungen sich auch so charakterisieren lassen, daß alle Transitivitätsklassen ($\neq \mathrm{E}$, E) von G und Γ unendlich sind. Gibt es unter diesen Transitivitätsklassen mindestens eine selbstin-

verse, so sprechen wir kurz über Fall a, sonst sprechen wir über Fall b. (Wir konnten die Existenz keiner dieser Fälle bekräftigen oder widerlegen.)

Fall a bedeutet, daß es ein Paar A $(\neq E)$, A $(\neq \mathsf{E})$ gibt, wofür eine der Gleichungen

$$(61) \qquad A^{\mathsf{A}} = A^{-1}, \quad \mathsf{A}^{A} = \mathsf{A}^{-1} \qquad (A \neq E,\ \mathsf{A} \neq \mathsf{E})$$

gilt. Nun besagt aber Hilfssatz 14 für diesen Fall, daß beide Gleichungen (61) einander gegenseitig bedingen. Hieraus folgt auch, daß im vorliegenden Fall a beide Gruppen G, Γ mindestens eine selbstinverse Transitivitätsklasse $(\neq E, \mathsf{E})$ enthalten. Ferner sieht man, daß (61) auch gleichbedeutend mit

$$(62) \qquad (A\mathsf{A})^2 = 1$$

ist (mit 1 bezeichnen wir hier das Einselement von $\mathfrak{G} = G\,\Gamma$). Da (62) unmöglich ist, wenn aus den Gleichungen $A = E$, $\mathsf{A} = \mathsf{E}$ nur die eine gilt, so ist Fall a gleichbedeutend damit, daß es in $\mathfrak{G} = G\,\Gamma$ Elemente zweiter Ordnung gibt. Ihre Zahl ist dann unendlich, denn zu jedem Element A bzw. A einer unendlichen Transitivitätsklasse gibt es ein A bzw. A mit (62), d. h. mit $O(A\mathsf{A}) = 2$. Genauer gilt auch, daß zu jedem A oder A höchstens nur ein A bzw. A gibt, so daß $O(A\mathsf{A}) = 2$ gilt, denn ist z. B. A vorgegeben, so existiert höchstens ein A mit (61_1). Wenn also z. B. A die Elemente aller selbstinversen Transitivitätsklassen $(\neq E)$ von G durchläuft, so durchläuft das nach (62) bestimmte A ebenfalls alle selbstinversen Transitivitätsklassen $(\neq \mathsf{E})$ von Γ, gleichzeitig durchläuft $A\mathsf{A}$ alle verschiedenen Elemente zweiter Ordnung von $\mathfrak{G} = G\,\Gamma$. Es ist leicht aus einem Element zweiter Ordnung $A\mathsf{A}$ weitere solche abzuleiten, so daß man seine Konjugierten bildet. So ist z. B. auch

$$B^{-1} A \mathsf{A} B = A B^{\mathsf{A}} B^{-1} \mathsf{A}^{B}$$

von zweiter Ordnung. Folglich sind alle $A B^{\mathsf{A}} B^{-1}$ $(B \in G)$ verschieden und sie gehören in selbstinverse Transitivitätsklassen. Das ist im wesentlichen alles, was wir über Fall a feststellen können.

Fall b scheint noch schwerer zugänglich zu sein. Wir müssen uns jetzt mit der ärmlichen Feststellung begnügen, daß dann wegen Haupthilfssatz 2 sowohl G als auch Γ mindestens zwei Transitivitätsklassen $(\neq E, \mathsf{E})$ enthält.

Man kann das Problem der Bestimmung aller Lösungen (2) von (4), (5), (6) (in beiden Fällen a, b) folgenderweise umformulieren. Man definiere zwei Funktionen f, g durch

$$(P^x)^{\Pi^y} = P^{f(x,y)}, \quad (\Pi^x)^{P^y} = \Pi^{g(x,y)}.$$

Dann lauten (4), (5), (6) so

$$f(x, 0) = x, \quad f(0, y) = 0, \qquad\qquad g(x, 0) = x, \quad g(0, y) = 0,$$

$$f(x, y + z) = f(f(x, y), z), \qquad\qquad g(x, y + z) = g(g(x, y), z),$$

$$f(x+y,z) = f(x,z) + f(y, g(z,x)), \qquad g(x+y,z) = g(x,z) + g(y, f(z,x)).$$

Diese Funktionalgleichungen scheinen noch schwerer handlich zu sein als die ursprüglichen (4), (5), (6).

LITERATURVERZEICHNIS

[1] G. Hajós, Über einfache und mehrfache Bedeckung des n-dimensionalen Raumes mit einem Würfelgitter, *Math. Zeitschrift*, **47** (1942), S. 427—467.

[2] L. Rédei, Anwendungen des schiefen Produktes in der Gruppentheorie, *Journal für reine und angewandte Math.*, **188.** (Unter Druck.)

[3] L. Rédei, Kurzer Beweis des gruppentheoretischen Satzes von Hajós, *Commentarii Math. Helvetici*, **23** (1949), S. 272—282.

[4] T. Szele, Neuer vereinfachter Beweis des gruppentheoretischen Satzes von Hajós, *Publicationes Math.* (Debrecen), **1** (1949), S. 56—62.

[5] J. Szép, Über die als Produkt zweier Untergruppen darstellbaren endlichen Gruppen, *Commentarii Math. Helvetici*, **22** (1948), S. 31—33.

[6] J. Szép, On the structure of groups which can be represented as the product of two subgroups, *Acta Sci. Math.*, **12** A (1950), S. 57—61.

[7] J. Szép and L. Rédei, On factorisable groups, *Acta. Sci. Math.*, **13** (1950), S. 235—238.

[8] J. Szép, On factorisable simple groups, *Acta Sci. Math.* (Unter Druck.)

[9] J. Szép, On factorisable, not simple groups, *Acta Sci. Math.*, **13** (1950), S. 239—241.

[10] J. Szép, On simple groups, *Publicationes Math.* (Debrecen), **1** (1949), S. 98.

[11] G. Zappa, Costruzione dei gruppi prodotto di due dati sottogruppi permutabili tra loro, *Atti Secondo Congresso Unione Mat. Italiana Bologna* 1940, S. 115—125.

(Eingegangen am 4. Mai 1950.)

К ТЕОРИИ ФАКТОРИЗУЕМЫХ ГРУПП, I

Л. РЭДЭИ (Сегед)

(Резюме)

Пусть G, Γ, $\mathfrak{G}$ обозначают три группы. Если

$$\mathfrak{G} = G\Gamma \tag{1}$$

подразумевая, что $\mathrm{A}\mathsf{A}$ ($\mathrm{A} \in G$, $\mathsf{A} \in \Gamma$) охватывает все различные элементы $\mathfrak{G}$, тогда частные группы G, Γ называем комплементарными факторами группы $\mathfrak{G}$, а саму группу $\mathfrak{G}$ факторизованной группой. Факторизуемые группы были исследованы Заппа [11][1] Сеп [5]—[10] и автором [2]. В основе вопроса лежит теорема Заппа:

Если для групп $G, \Gamma, \mathfrak{G}$ имеет место (1), тогда имеется два однозначно определенные функции

$$\mathrm{A}^{\mathsf{A}} \in G, \quad \mathsf{A}^{\mathrm{A}} (\in \Gamma) \qquad (\mathrm{A} \in G, \mathsf{A} \in \Gamma) \tag{2}$$

со свойством

$$\mathsf{A}\mathrm{A} = \mathrm{A}^{\mathsf{A}}\mathsf{A}^{\mathrm{A}} \tag{3}$$

и с свойствами

$$\mathrm{A}^{\mathsf{E}} = \mathrm{A}, \ \mathrm{E}^{\mathsf{A}} = \mathrm{E}, \ \mathsf{A}^{\mathrm{E}} = \mathsf{A}, \ \mathsf{E}^{\mathrm{A}} = \mathsf{E} \tag{4}$$

$$\mathrm{A}^{\mathsf{BC}} = (\mathrm{A}^{\mathsf{C}})^{\mathsf{B}}, \quad \mathsf{A}^{\mathrm{BC}} = (\mathsf{A}^{\mathrm{B}})^{\mathrm{C}}. \tag{5}$$

$$(\mathrm{AB})^{\mathsf{C}} = \mathrm{A}^{\mathsf{C}}\mathrm{B}^{\mathsf{C}^{\mathrm{A}}}, \quad (\mathsf{AB})^{\mathrm{C}} = \mathsf{A}^{\mathrm{C}^{\mathsf{B}}}\mathsf{B}^{\mathrm{C}}. \tag{6}$$

И наоборот, если даны, группы G, Γ о которых предположим, что имеют общий единичный элемент но других общих элементов не имеют, тогда любые решения (2), (4), (5), (6) с учетом (3) определяет одну группу $\mathfrak{G}$ со свойством (1), произведенную элементами G и Γ.

Согласно этому составление групп $\mathfrak{G}$ по(1) для данных групп G, Γ сведено к решению системы функциональных уравнений и поэтому эта задача является основной проблемой теории Заппа—Сеп. Еще до сего времени ни в одном случае не было дано конкретных решений. Автор ставит целью решение этой проблемы в случае с циклическими группами G, Γ, с предположением, что группа G бесконечна (случай двух конечных групп G, Γ будет разбираться в сообщении II.) — В случае конечной группы Γ, проблема получила полное решение. Если и Γ бесконечно, тогда автор мог решить только часть проблемы, неразрешенным отстался вопрос тех решений, для которых соотношения

$$A^{\mathsf{A}} \neq A, \mathsf{A}^{A} \neq \mathsf{A} \, (A \in G, A \neq E, \mathsf{A} \in \Gamma, \mathsf{A} \neq \mathsf{E}).$$

ьсегда имеют место. По всей вероятности, такого решенля нет. Один простой, но интересный пример: если G и Γ бесконечные циклические группы, P и Π производящие элементы этих групп, то функции

$$(P^i)^{\Pi^k} = P^{i(-1)^k}, \quad (\Pi^k)^{P^i} = \Pi^{k+2gi\bar{k}} \qquad (i, k = 0, \pm 1, \pm 2, \ldots)$$

дают одно решение где g является целым числом, далее $\bar{k} = 0$ или 1, $\bar{k} \equiv k$ (мод 2). Из групп $\mathfrak{G}$, относящимся к здесь приведенным решениям, все не изоморфные типы выявляются из двух случаев $g = 0, 1$.

[1] Скобки „[ ]" — указывают на список литературы, приложенный к немецкому тексту.

К ТЕОРИИ ПРЕДЕЛЬНЫХ ТЕОРЕМ ДЛЯ СУММ НЕЗАВИСИМЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

Член-корреспондент АЛЬФРЕД РЕНЬИ (Будапешт)

ВВЕДЕНИЕ

Пусть $\boldsymbol{M}$ — некоторое множество элементов $a, b, \ldots$ (в дальнейшем $\boldsymbol{M}$ будет называн основным пространством), $\boldsymbol{P}$ — борелевское поле подмножеств $A, B, \ldots$ пространства $\boldsymbol{M}$ и $\mu(A)$ — непрерывная мера *Lebesgue*-а. определенная для $A \varepsilon \boldsymbol{P}$; предположим $\boldsymbol{M} \varepsilon \boldsymbol{P}$ и $\mu(\boldsymbol{M}) = 1$. Измеримую относительно μ вещественную функцию $\xi = \xi(a)$ $(a \varepsilon \boldsymbol{M})$ называем случайной величиной, через $E(\xi < x)$ обозначим множество тех $a \varepsilon \boldsymbol{M}$ для которых $\xi(a) < x$ и положим $P[\xi < x] = \mu(E(\xi < x))$; т. е. $P[\xi < x]$ озвачает вероятность события $\xi < x$.

В настоящей статьи изучаются некоторые специальные последовательности случайных величин $\xi_n = \xi_n(a)$ $(n = 1, 2, 3, \ldots)$ каждая из которых принимает лишь конечное число различных значений[1] и обладающие свойством, что $\xi_n(a) = \xi_n(b)$ для всех $n = 1, 2, 3, \ldots$ не возможно. если $a \neq b$ (кроме быть может для точек a, b некоторого множества меры 0); такие последовательности случайных величим будем называть **максимальными**.

Пусть $\{\xi_n\}$ — максимальная последовательность независимых случайных величин, $\zeta_n = \xi_1 + \xi_2 + \ldots + \xi_n$, и положим $M(\zeta_n) = A_n$, $M((\zeta_n - A_n)^2) = B_n^2$, где $M(\xi)$ означает математическое ожидание величины ξ, т. е.

$$M(\xi) = \int_{M} \xi d\mu$$

и предположим, что закон распределения величины

$$\eta_n = \frac{\zeta_n - A_n}{B_n}$$

стремится для $n \to \infty$ к нормалному закону

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt \tag{1}$$

[1] От этого ограничения можно было бы освободиться.

В настоящей работе доказано, что если мера μ заменяется какой-либо другой мерой μ', которая абсолютно непрерывна относительно μ, то предельный закон распределения последовательности η_n остается неизменным (несмотря на то, что относительно меры μ' величины ξ_n вообще не являются независимыми и их законы распределения совсем изменяются). С другими словами, если $A_n(x) = E[\eta_n < x]$ из

(2) $$\lim_{n\to\infty} \mu(A_n(x)) = \Phi(x) \qquad (-\infty < x < +\infty)$$

следует

(3) $$\lim_{n\to\infty} \mu'(A_n(x)) = \Phi(x) \qquad (-\infty < x < +\infty)$$

если только μ' абсолютно непрерывная относительно μ.[2] Указанное свойство максимальных последовательностей случайных величин является следствием того, что последовательности множеств $A_n(x)$ имеют „перемещающиеся свойства", значит для всех $A \in P$ имеет место

(4) $$\lim_{n\to\infty} \{ \mu(A \, . \, A_n(x)) - \mu(A)\,\mu(A_n(x)) \} = 0.$$

Это свойство можно выразить также следующим образом: случайные величины η_n в пределе независимы от всякой случайной величины ξ; в самом деле, пусть $A = E(\xi < y)$, то из (4) получим

(5) $$\lim_{n\to\infty} \{ P[\eta_n < x; \xi < y] - P[\eta_n < x] . P[\xi < y] \} = 0.$$

Наиболее простой пример максимальной последовательности получим, если за $\boldsymbol{M}$ примем отрезок (0,1), за μ — обычную меру *Lebesgue*-а и за ξ_n — функции *Rademacher*-а, т. е. $\xi_n = R_n(t) = sg\,(\sin 2^n \pi t)$ $(0 \leq t \leq 1)$. В самом деле, значение t однозначно определяется значениями $R_n(t)$, $n = 1, 2, 3, \ldots$, так как

(6) $$t = \frac{1}{2} - \sum_{n=1}^{\infty} \frac{R_n(t)}{2^{n+1}}.$$

В этом случае наш результат может быть сформулирован следующим образом: Пусть $A_n(x)$ означает множество тех вещественных чисел t $(0 \leq t \leq 1)$ в разложении которых по двойчной системе разница между числом нулей и числом единицей среди первых n цифр не превосходит $\frac{x\sqrt{n}}{2}$, и пусть μ некоторая абсолютно непрерывная мера

(7) $$\mu(A) = \int_A \lambda(t)\,dt$$

[2] То же самое имеет место, если предельный закон распределения величин η_n отличается от нормального закона, но мы ограничимся рассмотрением этого наиболее важного случая.

где $\lambda(t)$ измеримая неотрицательная функция и $\int\limits_0^1 \lambda(t)\,dt = 1$, тогда имеем

$$\lim_{n\to\infty} \mu(A_n(x)) = \Phi(x). \tag{8}$$

Рассмотрим частный случай $\lambda(t) = \frac{1}{r} t^{\frac{1}{r}-1}$ $(r > 0)$: тогда (8) равносильно следующему утверждению: путь $A_n^{(r)}(x)$ означает множество тех вещественных чисел $t(0 \leq t \leq 1)$ для которых в разложении числа t^r по двойчной системе разница между числом нулей и числом единицей среди первых n цифр не превосходит $\frac{x\sqrt{n}}{2}$, тогда имеем

$$\lim_{n\to\infty} |A_n^{(r)}(x)| = \Phi(x) \tag{9}$$

где $|A_n^{(r)}(x)|$ означает обычную меру *Lebesgue*-а множества $A_n^{(r)}(x)$.

Заметим лишь что условие максимальности не является очень ограничительным, так как всякая последовательность независымых случайных величин может быть превращена в максималную отождествляя все точки в которых значения ξ_n совпадают для всех n, т. е. переходя к рассмотрению фактор — пространства пространства M по некоторому разбиению. Что касается теории измеримых разбиений пространства *Lebesgue*-а и вообще теории меры *Lebesgue*-а, мы будем пользоваться результатами В. А. Рохлина [1].

Обнаруженная устойчивость предельного распределения при изменении вероятностной меры свидетельствует о том, что предельные законы теории вероятностей не теряют силу, если обычние предположения (например предположение независимости) в некоторой степени нарушаются, как это хорошо известно из исследованиях С. Н. Бернштейна и [2] др. о слабо зависимых случайных величин. В самом деле, результаты настоящей работы могут быть истолкованы, как дающие обобщение предельных законов теорий вероятностей на некоторый класс слабо зависимых случайных величин.

§ 1. ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ И ОБОЗНАЧЕНИЯ

Пусть $\boldsymbol{M}$ — основное пространство. $\boldsymbol{P}$ борелевское поле подмножеств пространства $\boldsymbol{M}$, $\mu(A)$ — вполне аддитивная неотрицательная функция множеств, определенная на $\boldsymbol{P}$. Предположим что $\boldsymbol{M} \varepsilon \boldsymbol{P}$ и $\mu(\boldsymbol{M}) = 1$. далее что если $\mu(A) = 0$, и $B \varepsilon A$, то $B \varepsilon \boldsymbol{P}$ (и следовательно $\mu(B) = 0$). В этом случае $\boldsymbol{M}$ называется пространством с мерой. Элементы пространства $\boldsymbol{M}$ будем называть точками и обозначить малыми буквами, элеметы $\boldsymbol{P}$ называем измеримыми множествами и обозначим большими буквами. Предположим, что пространство $\boldsymbol{M}$ сепарабельно, т. е. что существует счетная система $\Delta = \{ D_n \}$ измеримых множеств, которая является базисом пространства $\boldsymbol{M}$, значит обладает следующими двумя свойствами:

Б. 1. Для любой пары точек a, b, $a \neq b$, пространства $\boldsymbol{M}$ (кроме быть может точек множества меры 0) сущетвует множество $D_n \varepsilon \mathcal{A}$ которое содержит только одну из точек, a, b.

Б. 2. Пусть $\boldsymbol{P}(\mathcal{A})$ наименьшее борелевское поле подмножеств пространства $\boldsymbol{M}$ которое содержит все множества D_n системы $\mathcal{A}$; тогда для всякого измеримого A существует содержащее A множество $B \varepsilon \boldsymbol{P}(\mathcal{A})$ которое мод. О тождественно с A т. е. $\mu(B-A)=0$.

Предположим дальше что $\boldsymbol{M}$ является полным относительно своего базиса, т. е для всякой последовательности патуральных чисел $n_1 < n_2 < \ldots < n_k < \ldots$ существует точка a так что $a \varepsilon D_{n_k}$ и $a \bar{\varepsilon} D_n$ если $n \neq n_k$ $(k=1, 2, \ldots)$. Из Б. 1. следует что точка a однозначно определена, т е. если положим $D_n^1 = D_n$ и $D_n^{-1} = \boldsymbol{M} - D_n$ то всякое множество $\prod_{n=1}^{\infty} D_n^{i_n}$ где $i_n = \pm 1$, одноточечно. Сепарабельное и полное относительно своего базиса пространство с мерой называется пространством *Lebsgue*-а. В далнейшем предположим что простарнство $\boldsymbol{M}$ с мерой μ является пространством *Lebsgue*-а, и кроме того, что мера μ непрерывна, т. е. что в $\boldsymbol{M}$ нет точек положительной меры. Известно, что пространство *Lebsgue*-а с непрерывной мерой по модулю 0 изоморфно единичному отрезку с обычной мерой *Lebesgue*-а. (См. [1.])

§ 2. МАКСИМАЛЬНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ НЕЗАВИСИМЫХ ЭЛЕМЕНТАРНЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

Пусть $\boldsymbol{M}$ пространство *Lebsgue*-а с непрерывной мерой μ; $\boldsymbol{M}$ может быть рассматриван как поле вероятностей в смысле А. Н. Колмогорова. Вещественную функцию $\xi = \xi(a)$, $(a \varepsilon \boldsymbol{M})$ определенную на $\boldsymbol{M}$ измеримую относительно μ будем называть случайной величиной. Случайную величину, которая принимает лишь конечное число различных значений, называем элементарной случайной величиной. Пусть ξ_n — последовательность элементарных случайных величин, положим $E_{nk} = E(\xi_n = x_{nk})$ $(k = 1, 2, \ldots, K_n$ и $n = 1, 2, \ldots)$: предположим далее что величины ξ_n независимы т. е. что

$$\mu\left(\prod_{i=1}^{j} E_{n_i k_i}\right) = \prod_{i=1}^{j} \mu(E_{n_i k_i}) \qquad (n_1 < n_2 < \ldots < n_j). \tag{2.1}$$

Введем следующее определение:

М. 1. Последовательность независимых элементарных случайных величин ξ_n называется максимальной если из $\xi_n(a) = \xi_n(b)$ для $n = 1, 2, 3, \ldots$ следует $a = b$ (кроме быть может для точек a, b некоторого множества меры 0).

Так как известно (см. [1]) что в случае пространства *Lebesgue*-а Б. 2. является следствием Б. 1., получим

М. 2. Последовательность независимых элементарных случайных величин (ξ_n) максимальна тогда и только тогда, если система $\Delta = \{E_{nk}\}$ где $E_{nk} = E(\xi_n = x_{nk})$ является базисом пространства $\boldsymbol{M}$.

Всякая последовательность независимых элементарных случайных величин может быть превращена в максимальную с заменой поля вероятностей. В самом деле пусть $D_{nk} = E_{n1} + E_{n2} + \ldots + E_{nk}$ $(k < K_n;\ n = 1, 2, \ldots)$, $D^{1}_{nk} = D_{nk}$ и $D^{-1}_{nk} = \boldsymbol{M} - D_{nk}$, и обозначим через C всякое непустое множество вида $\prod_{n=1}^{\infty}\prod_{k=1}^{K_n} D^{i_{nk}}_{nk}$ где $i_{nk} = \pm 1$: множества C не пересекаются и покрывают $\boldsymbol{M}$; таким образом мы получим измеримое разбиение пространства $\boldsymbol{M}$, которое обозначим через $\boldsymbol{R}$. Множества C могут быть рассматриваемы как точки нового пространства $\boldsymbol{N}$. Подмножество X множества $\boldsymbol{N}$ будем называть измеримым, если совокупность Y тех точек a, которые принадлежат к некоторому $C \in X$ измерима, и положим $\nu(X) = \mu(Y)$: пространство $\boldsymbol{N}$ с мерой ν называется фактор-пространством множества $\boldsymbol{M}$ по разбиению $\boldsymbol{R}$, и обозначается через $\boldsymbol{M}/\boldsymbol{R}$. Ввиду того что фактор-пространство пространства *Lebesgue*-а по некоторому его измеримому разбиению является также пространством *Lebesgue*-а (см. [1]) $\boldsymbol{N}$ с мерой ν будет пространством *Lebesgue*-а; легко видеть что мера ν непрерывна тогда и только тогда если положив

$$q_n = 1 - \operatorname*{Max}_{k \le K_n} P[E_{nk}], \text{ ряд } \sum_{n=1}^{\infty} q_n \text{ расходится.}$$

§ 3. ИНВАРИАНТНОСТЬ ПРЕДЕЛЬНОГО ЗАКОНА ПРИ ЗАМЕНЕ МЕРЫ

Прежде всего докажем следующую лемму:

Лемма: Пусть $\eta_n^{(1)}$ и $\eta_n^{(2)}$ независимые случайные величины $\eta_n = \eta_n^{(1)} + \eta_n^{(2)}$ для $n = 1, 2, \ldots$; пусть $F_n(x) = P(\eta_n < x)$, дальше $F_n^{(j)}(x) = P(\eta_n^{(j)} < x)$ для $j = 1, 2$. Если $\lim\limits_{n\to\infty} F_n(x) = F(x)$, $(-\infty < x < +\infty)$ где $F(x)$ непрерывная функция распределения и $\lim\limits_{n\to\infty} F_n^{(2)}(x) = 0$ если $x < 0$ и $\lim\limits_{n\to\infty} F_n^{(2)}(x) = 1$ если $x > 0$, то имеем также $\lim\limits_{n\to\infty} F_n^{(1)}(x) = F(x)$ $(-\infty < x < +\infty)$.

Доказательство леммы:

Пусть $f_n(t) = M(e^{i\eta_n t})$ характеристическая функция величины η_n, дальше $f_n^{(j)}(t) = M(e^{it\eta_n^{(j)}})$ характеристическая функция величины $\eta_n^{(j)}$ $(j = 1, 2)$ и наконец

$$(t) = \int_{-\infty}^{+\infty} e^{ixt}\,dF(x).$$

Тогда имеем

(3.1) $$\lim_{n\to\infty} f_n(t) = f(t)$$

и

(3.2) $$\lim_{n\to\infty} f_n^{(2)}(t) = 1$$

Так как $f_n(t) = f_n^{(1)}(t) f_n^{(2)}(t)$, из (3.1) и (3 2) следует что

(3.3) $$\lim_{n\to\infty} f_n^{(1)}(t) = f(t).$$

В силу известной теоремы о характеристических функциях (3.3) равносилна утверждению леммы.

Введем теперь некоторые обозначения: пусть $P_A(B)$ означает условную вероятность события B при условии A, т. е. положим $P_A(B) = \frac{P(AB)}{P(A)}$. Если $F(x) = P[\xi < x]$ — функция распределения случайной величины ξ, то обозначим через $F(x; A)$ условную функцию распределения величины при условии A, т. е. положим

(3.4) $$F(x; A) = P_A(\xi < x).$$

Докажем теперь следующую теорему:

ТЕОРЕМА 1. Пусть $\{\xi_n\}$ — максимальная последовательность независимых элементарных случайных величин, $\zeta_n = \xi_1 + \xi_2 + \ldots + \xi_n$, $M(\zeta_n) = A_n$ и $M[(\zeta_n - A_n)^2] = B_n^2$; предположим что $B_n \to \infty$. Пусть $\eta_n = \frac{\zeta_n - A_n}{B_n}$, $F_n(x) = P(\eta_n < x)$ и A некоторое событие, $\mu(A) > 0$. Если

(3.5) $$\lim_{n\to\infty} F_n(x) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt \qquad (-\infty < x < +\infty)$$

то имеем также

(3.6) $$\lim_{n\to\infty} F_n(x; A) = \Phi(x).$$

Доказательство Теоремы 1. Пусть $E_{nk} = E(\xi_n = x_{nk})$ $(k < K_n;$ $n = 1, 2, \ldots)$ и предположим вначале, что $A = \prod_{r=1}^{N} E_{rk_r}$. Тогда ζ_N постоянная на A, $\zeta_N(a) = Z_N$ для $a \in A$ и в силу независимости величин ξ_n получим для $n < N$.

(3.7) $$F_n(x; A) = P(\eta_{n\,N} < x)$$

где

(3.8) $$\eta_{n\,N} = \eta_n - \frac{(\eta_N \cdot B_N - Z_N)}{B_n}.$$

Пусть

(3.9) $$\vartheta_n = \frac{\eta_N B_N - Z_N}{B_n}.$$

тогда имеем

(3.10) $$\lim_{n\to\infty} P(\vartheta_n < x) = \begin{cases} 0 \text{ если } x < 0 \\ 1 \text{ если } x > 0 \end{cases}$$

так как $B_n \to \infty$ Применяя Лемму с $\eta_n^{(1)} = \eta_{nN}$ и $\eta_n^{(2)}$ ϑ_n, получим

(3,11) $$\lim_{n\to\infty} P(\eta_{n\cdot N} < x) = \Phi(x).$$

Таким образом Теорема 1. доказана в частном случае, если A имеет вид $\prod_{r=1}^{N} E_{rk_r}$; но так как имеем

(3.12) $$F_n(x; E_1 + E_2 + \ldots + E_s) = \sum_{j=1}^{S} \frac{\mu(E_j)}{\mu(E)} F_n(x_j; E_j),$$

если $E_iE_j = 0$ для $i \neq j$, сразу следует, что Теорема 1. верна также если A есть сумма конечного числа множеств вида $\prod_{r=1}^{N} E_{rk_r}$; из этого следует, что (3.6) имеет место, если A принадлежит к наименьшему борелевскому полю, содержащему все множества вида $\prod_{r=1}^{N} E_{rk_r}$; так как последовательност $\{\xi_n\}$ максимальна, это значит (см. М. 2 и Б 2) что (3 6) имеет место для всякого измеримого множества A, и таким образом Теорема 1. доказана.

Теперь перейдем к доказательству теоремы о замене меры:

ТЕОРЕМА 2. Пусть μ' относительно μ абсолютно непрерывная мера,

(3.13) $$\mu'(A) = \int_A \lambda(a)\, d\mu$$

где $\lambda(a) \geqq 0$ измеримая по μ функция, определенная на $\boldsymbol{M}$, и

(3.14) $$\int_M \lambda(a)\, d\mu = 1.$$

Тогда при условиях Теоремы 1., если положим $A_n(x) = E(\eta_n < x)$ имеем

(3.15) $$\lim_{n\to\infty} \mu'(A_n(x)) = \Phi(x).$$

Доказательство Теоремы 2.:

Предположим вначале, что $\lambda(a)$ „кусочно постоянная“ функция, другими словами, что $\lambda(a)$ принимает лишь конечное число различных значении. $\lambda(a) = \lambda_i$ если $a \varepsilon A_i$, $i = 1, 2, \ldots r$. Пусть $\mu(A_i) = \alpha_i$ тогда из (3.14) находим

(3.16) $$\sum_{=}^{r} \lambda_i \alpha_i = 1.$$

Очевидно имеем

$$\mu'(A_n(x)) = \sum_{i=1}^{r} \lambda_i \alpha_i F_n(x_i; A_i) \tag{3.17}$$

и поэтому применяя Теорему 1 и имея (3.16) в виду следует

$$\lim_{n\to\infty} \mu'(A_n(x)) = \Phi(x); \tag{3.18}$$

это значит, что (3.15) имеет место, если $\lambda(a)$ кусочно постоянна; из этого следует, что (3.15) верно, если $\lambda(a)$ ограниченная функция (так как всякая органиченная измеримая функция может быть как угодно точно приближена с помощью кусочно постоянных функций); оттого следует что (3.15) имеет место также, если $\lambda(a)$ произвольная измеримая функця (и (3.14) удовлетворена) таким образом Теорема 2. доказана.

§ 4. НЕКОТОРЫЕ СЛЕДСТВИЯ И ЗАМЕЧАНИЯ

ТЕОРЕМА 3. Пусть $b = f(a)$ измеримое отображение пространства $\boldsymbol{M}$ на самое себя, и рассмотрим меру μ', которая определена следующим образом:

$$\mu'(A) = \mu(f^{-1}[A]) \tag{4.1}$$

где $f^{-1}[A] = E(f(a) \varepsilon A)$; предположим что μ' абсолютно непрерывна относительно μ. При условиях Теоремы 1., если

$$\vartheta_n(a) = \eta_n(f(a)) \quad (a \varepsilon M) \tag{4.2}$$

и

$$G_n(x) = P(\vartheta_n < x) \tag{4.3}$$

следует

$$\lim_{n\to\infty} G_n(x) = \Phi(x). \tag{4.4}$$

Теорема 3. следует непосредственно из Теоремы 2. так как

$$G_n(x) = \mu'(E(\eta_n < x)). \tag{4.5}$$

Замечание о функциях *Rademacher*-а в введении (см. (9)) является частным случаем Теоремы 3.

ТЕОРЕМА 4. Пусть $g(t, u)$ функция двух вещественных переменных которая имеет для всякого фиксированного значения u однозначную обратную функцию $g_u^{-1}(x)$ $(g(g_u^{-1}(x), u) = x)$. Пусть ζ элементарная случайная величина, $F(x) = P(\zeta < x)$; рассмотрым случайные величины (4.6) $\vartheta_n = g(\eta_n, \zeta)$ и положим (4.7) $G_n(x) = P(\vartheta_n < x)$.

При условиях Теоремы 1 имеет место

$$\lim_{n\to\infty} G_n(x) = \int_{-\infty}^{+\infty} \Phi(g_u^{-1}(x))\, dF(u). \tag{4.8}$$

Теорема 4. является непосредственным следствием Теоремы 1.

Дальнейшие обобщения результатов настоящей работы и некоторые применения будут даны в другом статье. Заметим лишь что результаты настоящей работы связаны с некоторыми прежними исследованиями автора [3].

(Поступило 11/IV 1950 г.)

ЛИТЕРАТУРА

[1] *В. А. Рохлин,* Об основных понятиях теории меры. *Математический Сборник* 25/67/:1 (1949), стр 107—150.

[2] *S. Bernstein,* Sur l'extension du théorème limite du calcul des probabilités aux sommes des quantités dépendantes, *Math. Ann.*, **97** (1926), pp. 1—59.

[3] *A. Rényi,* Sur un théorème général de probabilités, *Annales de l'Institut Fourier*, **1** (1949), pp. 43—52.

CONTRIBUTIONS TO THE THEORY OF INDEPENDENT RANDOM VARIABLES

By ALFRÉD RÉNYI

(Summary)

Let μ denote a Lebesgue measure defined on the abstract space **M**; let A, B, ... denote measurable sets. The sequence of random variables (i. e. — measurable real functions defined on **M**) $\xi_n = \xi_n(a)$ $(a \in \mathbf{M})$ $(n = 1, 2, \ldots)$ is called *maximal* if $\xi_n(a) = \xi_n(b)$ holds for $n = 1, 2, \ldots$ only if $a = b$ (except for a and b belonging to a certain set of measure 0). If $\xi = \xi(a)$ denotes a random variable, let $M(\xi)$ denote the mean value of ξ, $E(\xi < x)$ the event $\xi < x$, $P(\xi < x) = \mu(E(\xi < x))$ the probability of the event $\xi < x$. Let $P_A(B)$ denote the conditional probability of the event B relative to the event A and let

$$F(x; A) = P_A(\xi < x)$$

denote the conditional distribution function of ξ relative to A. The following results are proved:

Theorem 1. Let $\{\xi_n\}$ denote a maximal sequence of independent random variables, each assuming only a finite number of different values; let us put $\zeta_n = \xi_1 + \xi_2 + \ldots + \xi_n$, $M(\zeta_n) = A_n$, $M((\zeta_n - A_n)^2) = B_n$ and suppose $B_n \to \infty$; let us put $\eta_n = \dfrac{\zeta_n - A_n}{B_n}$ and $F_n(x) = P(\eta_n < x)$ and let A denote any event having positive probability $(\mu(A) > 0)$. If

$$\lim_{n\to\infty} F_n(x) = \Phi(x) \qquad (-\infty < x < +\infty) \tag{1}$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt,$$

it follows that

(2) $$\lim_{n\to\infty} F_n\ (x;\ A) = \Phi\ (x)\ .$$

Theorem 2. If μ' denotes a Lebesgue measure on $\boldsymbol{M}$ which is absolutely continuous with respect to μ, and $A_n(x) = E\ (\eta_n < x)$, we have, under the conditions of Theorem 1,

(3) $$\lim_{n\to\infty} \mu'\ (A_n\ (x)) = \Phi\ (x)\ .$$

Theorem 3. Let $b = f\ (a)$ denote a measurable transformation of $\boldsymbol{M}$ onto itself and let us define the measure μ' as follows:

(4) $$\mu'\ (A) = \mu\ (f^{-1}\ (A))$$

where $f^{-1}\ (A) = E\ (f\ (a) \in A)$; let us suppose that μ' is absolutely continuous with respect to μ. Under the conditions of Theorem 1, if we put $\vartheta_n\ (a) = \eta_n\ (f\ (a))$ and $G_n\ (x) = P\ (\vartheta_n < x)$, it follows

(5) $$\lim_{n\to\infty} G_n\ (x) = \Phi\ (x).$$

Theorem 4. Let $g\ (t,\ u)$ denote a function of two real variables which has for every fixed u a unique inverse function $g_u^{-1}\ (x)$ $(g\ (g_u^{-1}\ (x)\ ,\ u) = x\ )$. Let ξ denote an arbitrary random variable, and consider the sequence of random variables $F\ (x) = P\ (\xi < x)$; let us put $G_n\ (x) = P\ (\vartheta_n < x)$. Under the conditions of Therem 1, we have

(6) $$\lim_{n\to\infty} G_n(x) = \int_{-\infty}^{+\infty} \Phi(g_u^{-1}\ (x))\ dF\ (u)\ .$$

Theorem 2 follows from Theorem 1, while Theorems 3 and 4 are simple consequences of Theorems 1 resp. 2. The proof of Theorem 1 makes use of some results of V. A. Rohlin on measure theory [1].

ON CAUCHY'S CONVERGENCE TEST

By

LÁSZLÓ KALMÁR (Szeged), corresponding member of the Academy

1. This paper contains two proofs, seeming to be never published, of Cauchy's convergence test, i. e. of the theorem to the effect that the condition

(*C*) to any positive ε there is a $\varrho = \varrho(\varepsilon)$ such that we have $|a_m - a_n| < \varepsilon$ for every $m, n > \varrho(\varepsilon)$

is necessary and sufficient for the convergence of the infinite sequence $a_1, a_2, \ldots, a_n, \ldots$ Both proofs might be useful for lecture purposes. The first of them has the advantage over the usual class-room proofs of being direct, i. e. not based on the Bolzano-Weierstrass theorem. The second proof furnishes an equivalent form of Cauchy's test allowing some simplifications in Cantor's theory of real numbers, as to be shown in a subsequent paper.

2. Cauchy's convergence test is proved in lectures on Calculus usually by means of the Bolzano-Weierstrass theorem, stating the existence of a convergent subsequence of every bounded infinite sequence. The latter theorem is usually proved by means of successive halving of the interval containing the sequence in question, choosing for the next halving in each case the half, or one of the halves, containing an infinity of terms of the sequence, and considering the common point of the intervals formed thus.

Now, the same idea, slightly modified, leads directly to (the sufficiency[1] of) Cauchy's convergence test. The modification consists in replacing the halves of an interval by two overlapping parts[2], e. g. by its left and right two-thirds. Indeed, let a_n be (the general term of) a sequence satisfying condition (*C*). We define b_n and c_n recursively as follows. Let $b_1 = a_{\varrho(1)+1} - 1$, $c_1 = a_{\varrho(1)+1} + 1$, and, given b_r and c_r, let $b'_r = \dfrac{2b_r + c_r}{3}$, $c'_r = \dfrac{b_r + 2c_r}{3}$ and

[1] The necessity of condition (*C*) for the convergence of the sequence a_n is obvious.

[2] The idea of this modification is due to Brouwer who applied, in several publications, overlapping intervals for the construction of numbers the mere existence of which has been proved by means of non-overlapping intervals.

define $b_{r+1} = b_r, c_{r+1} = c'_r$ if for all but a finite number of terms of the sequence a_n we have $b_r < a_n < c'_r$, and $b_{r+1} = b'_r, c_{r+1} = c_r$ in the opposite case. We prove by induction $b_r < a_n < c_r$ for every r and for every but a finite number of n. Indeed, this is true for $r = 1$ by (C), $\varepsilon = 1$. Suppose it to be true for an r and to be false for $r + 1$. Then, by definition of b_{r+1} and c_{r+1}, we should have $c'_r \leqq a_m < c_r$ and $b_r < a_n \leqq b'_r$ for an infinity of values of m and n, hence $a_m - a_n \geqq c'_r - b'_r$, contrarily to (C), $\varepsilon = c'_r - b'_r$.

Let a be the common point of the intervals $b_r \leqq x \leqq c_r$; then we have $|a_n - a| < c_r - b_r = 2\left(\frac{2}{3}\right)^{r-1}$ for all but a finite number of n, hence $a_n \to a$.

3. We next show that (C) is equivalent to the condition[3]

(C') any two subsequences[4] of a_n differ but in a 0-sequence (i. e. in a sequence converging to 0).

Hence, Cauchy's convergence test is equivalent to the theorem:

Condition (C') is necessary and sufficient for the convergence of the sequence a_n.

Indeed, suppose the sequence a_n satisfies condition (C). Let a_{k_n} and a_{l_n} be any two subsequences of a_n $(k_1 < k_2 < \ldots, l_1 < l_2 < \ldots)$. Then for any positive ε, we have $|a_{k_n} - a_{l_n}| < \varepsilon$ for $k_n, l_n > \varrho(\varepsilon)$, hence, on account of $k_n, l_n \geqq n$, for $n > \varrho(\varepsilon)$, i. e., we have $a_{k_n} - a_{l_n} \to 0$, thus, a_n satisfies condition (C') too. On the other hand, suppose the sequence a_n does not satisfy condition (C), i. e., for a positive ε_0, and for every positive integer ϱ, there are integers $m = \mu(\varrho)$ and $n = \nu(\varrho)$ such that $m, n > \varrho$ and $|a_m - a_n| \geqq \varepsilon_0$. Define $k_1 = l_1 = 1$ and, given k_r and l_r, let $k_{r+1} = \mu(\max(k_r, l_r))$ and $l_{r+1} = \nu(\max(k_r, l_r))$. Then we have $k_1 < k_2 < \ldots$, $l_1 < l_2 < \ldots$ and $|a_{k_n} - a_{l_n}| \geqq \varepsilon_0$ for every n. Hence $a_{k_n} - a_{l_n}$ does not converge to 0 and the sequence a_n does not satisfy condition (C').

4. Now, we prove the theorem, equivalent to Cauchy's convergence test, formulated in the preceding section. The necessity of condition (C') for the convergence of the sequence a_n being obvious, we have to prove its sufficiency only. First we prove that a sequence a_n satisfying condition (C') is necessarily bounded. Indeed, let a_n be a sequence unbounded from above. Define k_1 as the least k for which we have $a_k > a_1 + 1$ and, given k_r, define k_{r+1} as the least $k > k_r$ for which we have $a_k > a_{r+1} + 1$ (existing, for otherwise $a_{r+1} + 1$ would be an upper bound for the sequence $a_{k_r+1}, a_{k_r+2}, \ldots$,

[3] See also K. KNOPP, *Theorie und Anwendung der unendichen Reihen* (Berlin, 1931) pp. 89–90 (II. Hauptkriterium (3. Form)).

[4] "Subsequence" is meant in this paper so as to include the whole sequence too. Obviously, (C') is equivalent to the condition that a_n differs from any of its subsequences but in a 0-sequence.

and so, the whole sequence a_n would be bounded from above). Then, for $l_n = n$, we have $a_{k_n} > a_{l_n} + 1$ for every n, thus $a_{k_n} - a_{l_n}$ does not converge to 0 and the sequence a_n does not satisfy condition (C'). Similarly, we can show that a sequence unbounded from below does not satisfy condition (C').

Now, let a_n be a sequence satisfying condition (C') and a_{l_n} one of its convergent subsequences existing on account of the Bolzano-Weierstrass theorem. Let $a_{l_n} \to a$; by condition (C'), we have $a_n - a_{l_n} \to 0$; hence, $a_n \to a$.

5. In a lecture where Cauchy's convergence test is based on the proof given in the preceding two sections, it would be advisable, on reasons of style, to give a "qualitative" proof of the Bolzano—Weierstrass theorem too. The following such proof is due to the late Professor KÜRSCHÁK. It is an instance of a proof by cases of an unusual type[5].

Obviously, it suffices to prove that every infinite sequence has a monotonous subsequence; for if the sequence in question is bounded, the same holds for its subsequences too and so a monotonous subsequence of it is necessarily convergent. Now, let a_n be the sequence in question; define k_1 as the least k for which we have $a_k \leqq a_n$ for every n, and, given k_r, define k_{r+1} as the least $k > k_r$ for which we have $a_k \leqq a_n$ for every $n > k_r$. Of course, k_r may fail to exist; however, if it exists for every r, we have the increasing subsequence a_{k_n} of a_n. If, on the other hand, for a positive integer s, k_s does not exist, then the sequence $a_{k_s+1}, a_{k_s+2}, \ldots$ does not have a least term. In this case, define $l_1 = k_s + 1$ and, given l_r, define l_{r+1} as the least $l > l_r$ for which $a_l < a_{l_r}$ (existing, for otherwise a_{l_r} would be the least term of the sequence a_{l_r}, a_{l_r+1}, $a_{l_r+2}, \ldots$ and thus the sequence $a_{k_s+1}, a_{k_s+2}, \ldots$ formed of the former by subjoining a finite number of terms, would also have a least term[6]). Then a_{l_n} is a decreasing subsequence of a_n.

(Received 29 October 1949)

[5] This type of the proof by cases consists in giving first a proof method which does not work always and in giving another proof method for the case in which the first method does not work. The proofs for some particular cases, which I know from oral communication of Mr. BERNAYS, of the arithmetical lemma to which the consistency of Analysis has been reduced by ACKERMANN (see D. HILBERT, Die Grundlagen der Mathematik, *Abhandlungen aus dem math. Seminar der Hamburgischen Universität*, 6 (1928), pp. 65—85, especially the last paragraph beginning on p. 84), belong to the same type.

[6] As a matter of fact, as easily seen, a_{l_r} would be the least term of the sequence $a_{k_s+1}, a_{k_s+2}, \ldots$ too.

О КРИТЕРИИ СХОДИМОСТИ КОШИ

ЛАСЛО КАЛЬМАР (Сегед)

(Резюме)

Автор дает два новых доказательства критерия *Коши*, относящегося к сходимости бесконечных последовательностей. Первое доказательство аналогично обычному доказательству теоремы *Больцано—Вейерштрасса*, но применяет пересекающие интервалы; оно основывается на том, что если интервал содержит все члены бесконечной последовательности с исключением конечного числа членов, то или левая, или правая $^2/_3$ часть интервала имеет такие же свойства. Второе доказательство основывается на теореме *Больцано—Вейерштрасса* и дает критерий *Коши* в следующей форме: последовательность сходится тогда и только тогда, если от любой своей бесконечной под-последовательности она отличается только последовательностью сходящейся к нулю. Автор сообщает также одно доказательство *Кюрщака* теоремы *Больцано—Вейерштрасса*, которое основывается на том, что каждая последовательность имеет или возрастающую или убывающую под-последовательность.

ÜBER DIE DIOPHANTISCHE GLEICHUNG $x^3-1=2y^2$

Von

RICHARD OBLÁTH (Budapest)

(*Vorgelegt von* A. RÉNYI)

EULER[1] und T. NAGELL[2] haben die Unmöglichkeit der diophantischen Gleichung

$$x^3+1=y^2$$

für $x>2$ bewiesen. V. A. LEBESGUE[3] bewies, daß auch

$$x^3-1=y^2$$

(allgemeiner $x^n-1=y^2$) unlösbar ist. STÖRMER[4] hat diese Resultate ausgedehnt, indem er die Unmöglichkeit der Gleichung

$$x^2+1=2\,y^n$$

gezeigt hat. Prof. T. NAGELL bewies die Unmöglichkeit der Gleichungen

$$x^3\pm 1=y^n \qquad (n\geqq 3).$$

In diesen Gedankenkreis gehören auch die interessanten diophantischen Gleichungen

(I) $$x^3-1=2\,y^2$$

[1] L. EULER: Theorematum quorumdam arithmeticorum demonstrationes (1738), *Comm. Arithm. Coll.* **I** (1849), pp. 24–34, *Opera Omnia, Series prima* **II** (1915), pp. 38–58, speziell pp. 56–58, Theorema 10.

[2] T. NAGELL: Résultats nouveaux de l'analyse indéterminée I, *Norsk Matematisk Forenings Skrifter*, Série **I**, Nr. 8 (1922), § 10, pp. 11–13. Sur l'impossibilité de l'équation indéterminée $z^p+1=y^2$, ebenda Nr. 4 (1921), pp. 1–10, besonders pp. 1–2. Sur une équation diophantienne à deux indéterminées, *Det Kgl. Norske Videnskabers Selskabs Forhandlinger* **7**, Nr. 38, (1934), pp. 136–139.

[3] V. A. LEBESGUE: Sur l'impossibilité en nombres entiers de l'équation $x^m=y^2+1$. *Nouv. Ann. de Math.*, **9** (1850), pp. 178–181. Auch T. NAGELL, Analyse indéterminée de degré supérieur. *Mémorial des Sciences Mathématiques* Fasc. **39**, Paris (1929), p. 58.

[4] C. STÖRMER: Solution complète en nombres entiers de l'équation $m\,\mathrm{arctg}\,\frac{1}{x}+$ $+\,n\,\mathrm{arctg}\,\frac{1}{y}=k\frac{\pi}{4}$, *Bull. de la Soc. Math. de France*, **27** (1899), pp. 160–170, aber auch schon in 1895 in den Christiania Videnskabers Selskabs Skrifter. Auch NAGELL, a. a. O.[3], p. 58.

und

$$x^3 - 1 = 2\,y^n\,, \tag{II}$$

deren Unmöglichkeitsbeweis den Inhalt der gegenwärtigen Mitteilung bildet.

Im Laufe unserer Untersuchung werden wir auch anderen interessanten unmöglichen diophantischen Gleichungen begegnen. Zuerst beweisen wir den im Folgenden öfter anzuwendenden, an und für sich interessanten

Satz 1. *Die diophantische Gleichung*

$$x^2 + x + 1 = 3\,v^2 \tag{III}$$

mit ungeradem x hat außer den trivialen Lösungen $x = 1$, $v = 1$ keine weiteren Lösungen.

Beweis. Wenn (III) besteht, muß, da x als ungerade vorausgesetzt wurde, x die Form $x = 6a + 1$ und v die Form $v = 2b + 1$ besitzen. Dann ist also

$$3\,a\,(2\,a + 1) = 2\,b\,(b + 1)\,,$$

folglich $a = 2c$ und deshalb

$$3\,c\,(4\,c + 1) = b\,(b + 1)\,, \tag{1}$$

$c = 0$ führt zur trivialen Lösung $x = 1$, $v = 1$, die wir ausgeschlossen haben, so erhält man

$$b = k\,c \qquad \text{oder} \qquad b = k\,c - 1\,.$$

(1) ergibt im ersten Falle

$$(k^2 - 12)\,c = 3 - k\,,$$

was aber unmöglich ist, weil die zwei Seiten entgegengesetzte Vorzeichen besitzen. Im zweiten Falle bekommt man die ebenfalls unmögliche Gleichung

$$(k^2 - 12)\,c = 3 + k\,.$$

Wenn nämlich die linke Seite positiv ist, muß $k \geqq 4$ sein. Dies ist aber ausgeschlossen: die Unmöglichkeit von $k = 4$ springt in die Augen, wenn aber $k \geqq 5$ sein sollte, dann wäre

$$k^2 - 12 > k + 3\,,$$

daraus folgt aber $c < 1$, was doch unmöglich ist, weil c eine ganze Zahl ist. Q. e. d.

Diesen kurzen Beweis verdanke ich einer freundlichen Mitteilung des Herrn Prof. A. Rényi.

Satz 2. *Die diophantische Gleichung (I) ist* (außer der trivialen Lösung $x = 1$, $y = 0$) *unmöglich.*

Aus (I) würde

$$(x - 1)\,(x^2 + x + 1) = 2\,y^2$$

folgen. Denkbar wären die beiden Zerlegungen

$$A) \qquad x - 1 = 2u^2, \qquad x^2 + x + 1 = v^2,$$
$$B) \qquad x - 1 = 6u^2, \qquad x^2 + x + 1 = 3v^2,$$

denn es ist $(x^2 + x + 1) - (x - 1)^2 - 3(x - 1) = 3$ identisch, und darum kann der größte gemeinsame Teiler von $x^2 + x + 1$ und $x - 1$ nur gleich 1 oder 3 sein.

Der Fall *A)* ist sofort auszuschließen wegen

$$x^2 < x^2 + x + 1 < (x + 1)^2.$$

Im Falle *B)* zeigt die Gleichung $x = 6u^2 + 1$, daß x ungerade sein muß, die Bedingungen des Satzes 1 sind also erfüllt. Satz 1 zeigt, daß auch die Möglichkeit *B)* entfällt. Unser Satz ist also bewiesen.

Mein urspünglicher Beweis dieses Satzes beruhte auf einer descente infinie und war viel länger.

Als Verallgemeinerung des Satzes 2 beweisen wir den

SATZ 3. *Die diophantische Gleichung*

$$\text{(II)} \qquad x^3 - 1 = 2y^n \qquad (n \geqq 3)$$

ist (von den trivialen Lösungen $x = 1$, $y = 0$ und mit ungeradem n $x = -1$, $y = -1$, abgesehen) *in ganzen Zahlen unlösbar.*

Die beiden im vorausgegangenen Beweise benützten Zerlegungen

$$A) \qquad x - 1 = 2u^n, \qquad x^2 + x + 1 = v^n,$$
$$B) \qquad x - 1 = 2 \cdot 3^{n-1} u^n, \qquad x^2 + x + 1 = 3v^n$$

sind auch jetzt zu unterscheiden.

Die Unmöglichkeit der Gleichung

$$x^2 + x + 1 = 3v^n$$

mit den trivialen Ausnahmen ($x = 1$, $x = -2$, $v = 1$) wurde von Herrn T. NAGELL[5] für jeden Exponenten $n \geqq 3$ bewiesen, während die Unmöglichkeit von

$$\text{(2)} \qquad x^2 + x + 1 = v^n$$

(mit den Ausnahmen $x = 0$, $x = -1$, $v = 1$; $n = 3$, $x = 18$, $x = -19$, $v = 7$) in Anlehnung an ein NAGELLsches[5] Resultat von Herrn Prof. W. LJUNGGREN[6] für jeden Exponenten $n \geqq 3$ bewiesen wurde. Nun haben wir durch

[5] T. NAGELL, Des équations indéterminées $x^2 + x + 1 = y^n$ et $x^2 + x + 1 = 3y^n$, *Norsk Matematisk Forenings Skrifter*, Serie I, Nr. 2 (1921), pp. 1—14, besonders § 3, pp. 12—14.

[6] W. LJUNGGREN: Einige Bemerkungen über die Darstellung ganzer Zahlen durch binäre kubische Formen mit positiver Diskriminante, *Acta Math.*, 75 (1942), pp. 1—21, besonders p. 11, die Anmerkung.

Satz 2 die Ausnahmefälle erledigt, die Einsetzung der zulässigen Lösungen der Gleichung (2) in *A)* zeigt, daß sie keine, bzw. nur die triviale Lösung der Gleichung (II) ergeben, womit unser Satz als bewiesen erscheint.

Es wäre naheliegend, daran zu denken, daß die sehr analoge Gleichung

$$x^3 + 1 = 2\,y^2$$

ebenfalls unmöglich ist. Diese Gleichung hat aber außer den trivialen $x = -1$, $y = 0$; $x = 1$, $y = +1$ zwei nichttriviale Lösungen:[7] $x = 23$, $y = \pm 78$, so daß unser Satz auf diese Gleichung sicher nicht in unveränderter Weise ausgedehnt werden kann. Infolge des THUE—SIEGELschen, bzw. des LANDAU—OSTROWSKYschen Satzes hat sie jedenfalls bloß eine endliche Anzahl von Lösungen.

Wir zeigen aber, und zwar als weitere Anwendung des Satzes 1 den

SATZ 4. *Die diophantische Gleichung*

$$3\,x^2 + 1 = z^3 \tag{IV}$$

ist (von der trivialen Lösung $x = 0$, $z = 1$ abgesehen) *in ganzen Zahlen unmöglich.*

Beweis. (IV) ist identisch mit

$$z^3 - 1 = (z - 1)(z^2 + z + 1) = 3\,x^2.$$

Die beiden denkbaren Zerlegungen sind

$$\begin{array}{lll} A) & z - 1 = 3\,u^2, & z^2 + z + 1 = v^2, \\ B) & z - 1 = u^2, & z^2 + z + 1 = 3\,v^2. \end{array}$$

A) entfällt, denn aus $z \equiv 1 \bmod 3$ folgt zwar $z^2 + z + 1 \equiv 0 \bmod 3$, der Ausdruck links ist aber, wie bekannt, bloß durch 3, nicht durch 3^2 teilbar. Im Falle *B)* kann z wegen Satz 1 (außer dem trivialen Fall) nicht ungerade sein. z sei also gerade. Aus (IV) folgt, daß dann x ungerade sein muß, es ist also

$$3\,x^2 \equiv 3 \bmod 8,$$

während doch aus (IV)

$$3\,x^2 \equiv -1 \bmod 8$$

entspringt. Der Widerspruch beweist unseren Satz.

(Eingegangen am 15. Mai 1950.)

[7] AUBRY und FAUQUEMBERGUE: Sphinx-Oedipe, 6 (1911), pp. 103—104, und 8 (1913), pp. 170—171. Eine eventuelle Lösung > 23 besteht aus mindestens 256 Ziffern. S. auch L. E. DICKSON: History of the Theory of Numbers **II**, New York (1934), p. 630.

ДИОФАНТОВОЕ УРАВНЕНИЕ $x^3 - 1 = 2y^2$

РИХАРД ОБЛАТ (Будапешт)

(Резюме)

В работе показано невозможность нескольких диофантовых уравнений (за исключением тривиальных решений) напр.

$$x^3 - 1 = 2y^2$$
$$x^3 - 1 = 2y^n \qquad (n \geqq 3)$$
$$3x^3 + 1 = \text{[illegible]}^3$$

Кроме нескольких теорем *Нагеля* и *Люнгрена*, основным средством доказательства служит то, что диофантовое уравнение

$$x^2 + x + 1 = 3y^2,$$

за исключением тривиального решения $x = 1$, с нечетным x невозможно.

THE EXTENSION OF PARTIALLY ORDERED GROUPS

By

L. FUCHS (Budapest)

(*Presented by* G. HAJÓS)

1. The extension theory of algebraic systems is a young part of the algebra. At first O. SCHREIER[1] has proposed and solved the problem of group extensions: given two groups N and F, find all possible groups G such that N is an invariant subgroup of G and G/N is isomorphic to F. Analogous problems on some other systems have been discussed by other authors,[2] but it seems to me that the extension problem for partially ordered groups has not yet been studied in general.[3] The purpose of the present paper is to give a general discussion of the extension of partially ordered groups.

2. First we recall some definitions needed throughout the paper.

A group G, written additively, is said to be a *partially ordered group* if an order relation $\geqq$ is defined for some pairs of its elements satisfying the following postulates:[4]

(I) *reflexive law:* $x \geqq x$ for all x in G.

(II) *antisymmetric law:* $x \geqq y$ and $y \geqq x$ imply $x = y$,

(III) *transitive law:* $x \geqq y$ and $y \geqq z$ imply $x \geqq z$,

(IV) *homogeneity law:* $x \geqq y$ implies $u + x + v \geqq u + y + v$ for all u, v in G.

Sometimes it is useful to add a fifth postulate:

(V) *the MS-property* (MOORE—SMITH): for every x, y in G there exists a third element z in G such that $z \geqq x$ and $z \geqq y$.

[1] O. SCHREIER, Über die Erweiterung von Gruppen, I: *Monatshefte Math. Phys.*, **34** (1926), pp. 165—180; II: *Hamburger Abh.*, **4** (1926), pp. 321—346.

[2] For example, the extension theory for rings was given by C. J. EVERETT, JR, An extension theory for rings, *Amer. J. Math.*, **64** (1942), pp. 363—370. Independently of EVERETT, T. SZELE discovered the same extension theory.

[3] In a very special case G. BIRKHOFF has given a method for constructing certain, but not all extensions. See his paper: Lattice-ordered groups, *Annals Math.*, **43** (1942), pp. 298—331.

[4] See e. g. C. J. EVERETT and S. ULAM, On ordered groups, *Trans. Amer. Math. Soc.*, **57** (1945), pp. 208—216.

x and y are called *incomparable*, in sign $x \| y$, if neither $x \geqq y$ nor $y \geqq x$. By G^+ we denote the set of all *positive* elements of G, i. e., those satisfying $x \geqq 0$ where 0 is the group identity. Homogeneity implies that $x \geqq y$ if and only if $x - y$, or equivalently, $-y + x$ is positive, and that the set G^+ is invariant under all inner automorphisms of G.

If a partially ordered group G is at the same time a lattice under the same order relation $\geqq$ as inclusion relation, G is termed a *lattice-ordered group*. A *linearly ordered* group ($x \| y$ is impossible) is always lattice-ordered.

A group isomorphism Θ between two partially ordered groups G and H is said to be an *order-isomorphism* (o-isomorphism), if it preserves order in both directions, i. e., $x \geqq y$ in G if and only if $\Theta(x) \geqq \Theta(y)$ in H. If C is an invariant subgroup of G which is at the same time convex in the sense that $a \geqq x \geqq b$ (a, b in C) implies $x \in C$, then in the factor-group G/C a natural order is induced by the order in G. This is defined by putting $C + x \geqq C + y$ if and only if for some representatives $u \in C + x$, $v \in C + y$ a relation $u \geqq v$ holds. The fulfilment of (I)—(IV) in G/C is an immediate consequence of the convexity of G. Whenever we are speaking of G/C as a partially ordered group, we always mean G/C with the order induced by the order of G.[5]

3. If the partially ordered group G contains C as an invariant convex subgroup and the factor-group G/C is o-isomorphic to the partially ordered group F, then G will be called a partially ordered extension of C by F, briefly a $C-F$-extension. Given two partially ordered groups $C = \{0, a, b, \ldots\}$ and $F = \{o, \varrho, \sigma, \ldots\}$, our problem consists in finding all possible $C-F$-extensions.

Assuming the problem solved, let $\{s_\varrho\}$ be a complete representation system for the extension group G modulo C, so that under the o-isomorphism $G/C \longleftrightarrow F$ we have $s_\varrho \longleftrightarrow \varrho$. We may always normalize the representation system by assuming 0 to be chosen for the representative of the null-class mod C, that is,

$$s_0 = 0. \tag{1}$$

With this representation system $\{s_\varrho\}$ each element x of G may be written uniquely in the form

$$x = s_\varrho + a \quad \text{with} \quad \varrho \in F,\ a \in C.$$

A factor-set $\{f_{\varrho\,\sigma}\}$ is defined by the addition rule

$$s_\varrho + s_\sigma = s_{\varrho+\sigma} + f_{\varrho\,\sigma} \quad (f_{\varrho\,\sigma} \in C) \quad \textit{for all} \quad \varrho, \sigma \in F. \tag{2}$$

By our normality assumption (1) we clearly have

$$f_{0\,0} = f_{\varrho\,0} = f_{0\,\varrho} = 0 \quad \textit{for all} \quad \varrho \in F. \tag{3}$$

[5] For a detailed discussion see my paper: On partially ordered groups, *Proc. Kon. Nederl. Akad. v. Wetensch.*, **53** (1950), pp. 828—834.

Each s_ϱ defines an o-automorphism Θ_ϱ in C given by

$$a\,\Theta_\varrho = -s_\varrho + a + s_\varrho,$$

in particular, by (1),

(4) $$a\,\Theta_0 = a \quad \textit{for all} \quad a \in C.$$

Two o-automorphisms Θ_ϱ and Θ_σ may be composed by the rule

$$a\,(\Theta_\varrho\,\Theta_\sigma) = (a\,\Theta_\varrho)\,\Theta_\sigma = -s_\sigma - s_\varrho + a + s_\varrho + s_\sigma = \\ = -f_{\varrho\sigma} - s_{\varrho+\sigma} + a + s_{\varrho+\sigma} + f_{\varrho\sigma},$$

that is,

(5) $$a(\Theta_\varrho\,\Theta_\sigma) = -f_{\varrho\sigma} + a\,\Theta_{\varrho+\sigma} + f_{\varrho\sigma}.$$

The associative law in G implies

(6) $$f_{\tau+\sigma,\tau} + f_{\varrho\sigma}\,\Theta_\tau = f_{\varrho,\sigma+\tau} + f_{\sigma\tau}.$$

We have so far considered merely the group operation in G. In order to describe the order relation in G, we have to define sets $P_{\varrho\sigma}$, which may be called *order-sets*, consisting of all $a \in C$ satisfying $s_\varrho + a \geqq s_\sigma$. These sets $P_{\varrho\sigma}$ will in fact characterize the order in G, for, clearly, $s_\varrho + a \geqq s_\sigma + b$ if and only if $a - b \in P_{\varrho\sigma}$. For these sets we establish the following properties.

A relation $s_\varrho + a \geqq s_\sigma + b$ in G implies, by definition, a relation $\varrho \geqq \sigma$ in F, and if $\varrho = \sigma$, then $a \geqq b$. Therefore,

(7) $P_{\varrho\sigma}$ *is nonvacuous if and only if* $\varrho \geqq \sigma$; *further* $P_{\varrho\varrho} = C^+$ *for all* $\varrho \in F$.

The reflexive and antisymmetric laws for the order in G are now readily checked.

By the transitive law, $s_\varrho + a \geqq s_\sigma + b \geqq s_\tau + c$, that is, $a - b \in P_{\varrho\sigma}$ and $b - c \in P_{\sigma\tau}$ imply $s_\varrho + a \geqq s_\tau + c$, that is, $a - c \in P_{\varrho\tau}$. Hence it follows

(8) $$P_{\varrho\sigma} + P_{\sigma\tau} \subset P_{\varrho\tau} \quad \textit{for all} \quad \varrho \geqq \sigma \geqq \tau.$$

Left-homogeneity implies that $s_\varrho + a \geqq s_\sigma$ is equivalent to $c + s_\varrho + a \geqq c + s_\sigma$, or $s_\varrho + c\,\Theta_\varrho + a \geqq s_\sigma + c\,\Theta_\sigma$ for each $c \in C$, hence

(9) $$c\,\Theta_\varrho + P_{\varrho\sigma} - c\,\Theta_\sigma = P_{\varrho\sigma} \quad \textit{for all} \quad c \in C \quad \textit{and} \quad \varrho \geqq \sigma.$$

Again from left-homogeneity we conclude to the equivalence of $s_\varrho + a \geqq s_\sigma$ with $s_\tau + s_\varrho + a \geqq s_\tau + s_\sigma$, or with $s_{\tau+\varrho} + f_{\tau\varrho} + a \geqq$ $\geqq s_{\tau+\sigma} + f_{\tau\sigma}$ for every $\tau \in F$, whence we get

(10) $$f_{\tau\sigma} + P_{\varrho\sigma} - f_{\tau\sigma} = P_{\varrho+\tau,\sigma+\tau} \quad \textit{for all} \quad \tau \in F \quad \textit{and} \quad \varrho \geqq \sigma.$$

Right-homogeneity asserts that from $s_\varrho + a \geqq s_\sigma$ it follows $s_\varrho + a +$ $+ s_\tau \geqq s_\sigma + s_\tau$ and hence $s_{\varrho+\tau} + f_{\varrho\tau} + a\,\Theta_\tau \geqq s_{\sigma+\tau} + f_{\sigma\tau}$, and vice versa. Thus we obtain

(11) $$f_{\varrho\tau} + P_{\varrho\sigma}\,\Theta_\tau - f_{\sigma\tau} = P_{\varrho+\tau,\sigma+\tau} \quad \textit{for all} \quad \tau \in F \quad \textit{and} \quad \varrho \geqq \sigma,$$

where, evidently, $P_{\varrho\sigma}\Theta_\tau$ denotes the totality of elements of the form $x\Theta_\tau$ with $x \in P_{\varrho\sigma}$.

A system $\{f_{\varrho\sigma}, \Theta_\varrho, P_{\varrho\sigma}\}$ where $f_{\varrho\sigma} \in C$, $\Theta_\varrho \in A(C)$,[6] $P_{\varrho\sigma} \subset C$ will be called a $C-F$*-extension-set* if it satisfies (3)—(11). Thus every extension G with a representation system $\{s_\varrho\}$ defines some $C-F$-extension-set.

Conversely, if given two partially ordered groups C, F, any $C-F$-extension-set $\{f_{\varrho\sigma}, \Theta_\varrho, P_{\varrho\sigma}\}$ defines an extension. For, the system consisting of all symbols (ϱ, a) with $\varrho \in F$, $a \in C$ under the rules

(α) *equality:* $(\varrho, a) = (\sigma, b)$ *if and only if* $\varrho = \sigma$ *and* $a = b$,

(β) *composition:* $(\varrho, a) + (\sigma, b) = (\varrho + \sigma,\ f_{\varrho\sigma} + a\,\Theta_\sigma + b)$,

(γ) *order relation:* $(\varrho, a) \geqq (\sigma, b)$ *if and only if* $a - b \in P_{\varrho\sigma}$

becomes a partially ordered group H. In fact, conditions (3)—(6) imply that H is a group (SCHREIER) and conditions (7)—(11) ensure that H obeys the laws (I)—(IV) for partial order.[7] Since H has a convex invariant subgroup C^* consisting of all elements of the form $(0, a)$ o-isomorphic to C under the mapping $(0, a) \leftrightarrow a$,[8] and the factor-group H/C is o-isomorphic to F, our stated problem is solved if we know all $C-F$-extension-sets.

It can happen that two or more extension-sets define the same extension. Indeed, if we use another representation system $\{t_\varrho\}$ instead of $\{s_\varrho\}$, connected by the relations $t_\varrho = s_\varrho + c_\varrho$ $(c_\varrho \in C)$, then the $C-F$-extension-set $\{f'_{\varrho\sigma}, \Theta'_\varrho, P'_{\varrho\sigma}\}$ belonging to the system $\{t_\varrho\}$ will define the same extension. The new extension-set satisfies the relations

(a) $c_0 = 0$,

(b) $f'_{\varrho\sigma} = -c_{\varrho+\sigma} + f_{\varrho\sigma} + c_\varrho\,\Theta_\sigma + c_\sigma$,

(c) $a\Theta'_\varrho = -c_\varrho + a\,\Theta_\varrho + c_\varrho$,

(d) $P'_{\varrho\sigma} = -c_\varrho + P_{\varrho\sigma} + c_\varrho$,

which are readily obtained from the definitions by direct computations.

We shall call two $C-F$-extension-sets $\{f_{\varrho\sigma}, \Theta_\varrho, P_{\varrho\sigma}\}$ and $\{f'_{\varrho\sigma}, \Theta'_\varrho, P'_{\varrho\sigma}\}$ *equivalent,* if they are connected by elements c_ϱ of C satisfying (a)—(d). This is actually an equivalence relation in the usual sense. By definition, extension-sets defined by the same extension belong to the same class of equivalence.

The converse is also true: two equivalent extension-sets define abstractly the same extension; in other words, the extensions H_1 and H_2 defined by equivalent $C-F$-extension-sets are o-isomorphic under the mapping $(\varrho, a) \leftrightarrow (\varrho, c_\varrho + a)$ which leaves the elements of C and the cosets of F fixed. (For the proof, which is easy to perform, use (α)—(γ) as well as (a)—(d).)

[6] $A(C)$ denotes the group of all o-automorphisms of C.

[7] We observe that if one assumes the MS-property both in C and in F, then the extension group will also have the MS-property. The proof is immediate.

[8] Henceforth we shall identify C^* with C.

Hence there is a one-one correspondence between the $C-F$-extensions and the classes of equivalent $C-F$-extension-sets. We thus conclude

THEOREM 1. *Every $C-F$-extension defines a class of equivalent $C-F$-extension-sets and conversely, every class of equivalent $C-F$-extension-sets defines a $C-F$-extension which is unique up to an o-isomorphism leaving the elements of C and the cosets of F fixed.*

By this theorem the extension problem for partially ordered groups is completely solved.

4. Let C and F be two partially ordered groups. In the hierarchy of all possible $C-F$-extensions we shall point out a certain type of particular interest.

One may ask whether to any factor-set $\{f_{\varrho\,\sigma}\}$ and any system of o-automorphisms $\{\Theta_\varrho\}$ satisfying (3)—(6) there may be found a system of order-sets $\{P_{\varrho\,\sigma}\}$ completing them to a $C-F$-extension-set. This question will be answered in the affirmative by putting

$$P_{\varrho\,\sigma} = \begin{cases} C \text{ if } \varrho > \sigma, \\ C^+ \text{ if } \varrho = \sigma, \\ \text{empty otherwise.} \end{cases}$$

As is readily seen, under this definition all conditions (7)—(11) are necessarily fulfilled. Accordingly, we have in the extension group H^*:

$$(\varrho, a) \geqq (\sigma, b) \text{ if and only if } \varrho > \sigma, \text{ or } \varrho = \sigma \text{ and } a \geqq b;$$

therefore, it is natural to say that H^* is a *lexicographic $C-F$-extension.*

Let H^* be a fixed lexicographic $C-F$-extension and H any other one with the same factor-set $\{f_{\varrho\,\sigma}\}$ and the same set $\{\Theta_\varrho\}$ of o-automorphisms. The mapping $(\varrho, a) \rightarrow (\varrho, a)$ of H onto H^* is by (β) an isomorphism in the pure group-theoretic sense. Moreover, this group-isomorphism preserves order; however, its inverse mapping does not do it necessarily. Hence it is a tautomorphism in the sense defined by the author[9] and thus we arrive at

THEOREM 2. *If $\{f_{\varrho\,\sigma}\}$ and $\{\Theta_\varrho\}$ satisfy* (3)—(6), *the lexicographic extension always exists. Any $C-F$-extension H with the extension-set $\{f_{\varrho\,\sigma}, \Theta_\varrho, P_{\varrho\,\sigma}\}$ is tautomorphic to the lexicographic $C-F$-extension H^* with the same factor-set $\{f_{\varrho\,\sigma}\}$ and the same set $\{\Theta_\varrho\}$ of o-automorphism.*

5. There is an analogue of the direct sum extension of groups.

Let $f_{\varrho\,\sigma} = 0$ for all $\varrho, \sigma \in F$; $a\,\Theta_\varrho = a$ for all $a \in C$ and all $\varrho \in F$; further let $P_{\varrho\,\sigma} = C^+$ whenever $\varrho = \sigma$. Then the extension defined by this extension-set always exists; it is called the *direct sum* of the partially ordered groups C and F.

Of course, this is not the only extension-set by which the extension is direct, for any other equivalent extension-set defines the same extension.

[9] See loc. cit.[5]

From (a)—(d) in 3 it is evident that an extension with the extension-set $\{f_{\varrho\,\sigma}, \Theta_\varrho, P_{\varrho\,\sigma}\}$ is direct if and only if there are elements c_ϱ in C such that (a′) $c_0 = 0$, (b′) $f_{\varrho\,\sigma} = c_{\varrho+\sigma} - c_\varrho - c_\sigma$, (c′) $a\Theta_\varrho = c_\varrho + a - c_\varrho$, (d′) $P_{\varrho\,\sigma} = c_\varrho + C^+ - c_\sigma$ for $\varrho \geqq \sigma$.

The order definition of the direct sum extension:

$$(\varrho, a) \geqq (\sigma, b) \text{ if and only if } \varrho \geqq \sigma \text{ and } a \geqq b,$$

is so simple that one naturally arises the question as to whether there exist even other $C-F$-extensions with the same order definition. Otherwise expressed, we seek the $C-F$-extension-sets with the order-sets $P_{\varrho\,\sigma} = C^+$ for $\varrho \geqq \sigma$. We shall show that if we assume the MS-property in the factor-group F, then with this definition of order no extension exists except the direct one.

Conditions (9)—(11) with $P_{\varrho\,\sigma} = C^+$ $(\varrho \geqq \sigma)$ in turn imply

$$c\,\Theta_\varrho = c\,\Theta_\sigma \text{ for } \varrho \geqq \sigma \text{ and all } c \in C,$$
$$f_{\tau\,\varrho} = f_{\tau\,\sigma} \text{ and } f_{\varrho\,\tau} = f_{\sigma\,\tau} \text{ for } \varrho \geqq \sigma \text{ and all } \tau \in F.$$

Hence, using the MS-property in F, we conclude that $c\,\Theta_\varrho = c\,\Theta_0 = c$, i. e., Θ_ϱ is the identity mapping for all $\varrho \in F$, and $f_{\varrho\,\sigma} = f_{\varrho\,0} = 0$ for all $\varrho, \sigma \in F$, that is, the extension is direct, as stated.

6. We shall now give conditions under which the extension group becomes a lattice-ordered group. $\wedge$ and $\vee$ will denote the lattice operations meet and join, respectively.

Let G be a lattice-ordered group and assume that the convex invariant subgroup C is closed under the lattice operations. Then, whenever $\varrho \geqq \varphi$ as well as $\sigma \geqq \varphi$, the order-sets $P_{\varrho\,\varphi}$ and $P_{\sigma\,\varphi}$ always have a common element f; for this f we have $s_\varphi - f \leqq s_\varrho \wedge s_\sigma$. Therefore $s_\varrho \wedge s_\sigma = s_\tau + c$ implies $\tau \geqq \varphi$, consequently, τ is the greatest lower bound for ϱ and σ, and hence $\tau = \varrho \wedge \sigma$ exists. Similarly, $\varrho \vee \sigma$ also exists. Thus the factor-group has to be lattice-ordered.

Let $(s_\varrho + a) \wedge (s_\sigma + b) = s_{\varrho \wedge \sigma} + c$. The two inequalities $s_\varphi + f \leqq s_\varrho + a$ and $s_\varphi + f \leqq s_\sigma + b$ are together equivalent to the single one $s_\varphi + f \leqq s_{\varrho \wedge \sigma} + c$, whence we get ($\cap$ denotes set-theoretic intersection):

$$(12) \qquad (-a + P_{\varrho\,\varphi}) \cap (-b + P_{\sigma\,\varphi}) = -c + P_{\varrho \wedge \sigma, \varphi} \quad \textit{if} \quad \varphi \leqq \varrho \wedge \sigma,$$

where $-c$ *is the unique least element of the set* $(-a + P_{\varrho, \varrho \wedge \sigma}) \cap \cap (-b + P_{\sigma, \varrho \wedge \sigma})$.

Similarly,

$$(13) \qquad (P_{\varphi, \varrho} + a) \cap (P_{\varphi, \sigma} + b) = P_{\varphi, \varrho \vee \sigma} + c \quad \textit{if} \quad \varphi \geqq \varrho \vee \sigma,$$

where c *is the unique least element of the set* $(P_{\varrho \vee \sigma, \varrho} + a) \cap (P_{\varrho \vee \sigma, \sigma} + b)$.

Now if H is any $C-F$-extension, where both C and F are lattice-ordered groups and the extension-set satisfies in addition to (3)—(11) also (12) and (13),

the extension group H becomes a lattice-ordered group, as the reader may readily verify for himself.

We finally observe that the extension group is linearly ordered if and only if both C and F are linearly ordered and the extension is lexicographic.

(Received 5 April 1950)

РАСШИРЕНИЕ ЧАСТИЧНО УПОРЯДОЧЕННЫХ ГРУПП

Л. ФУКС (Будапешт)

(Резюме)

Автор, в соответствии с расширением групп по Шрейеру, дает расширение для частично упорядоченных групп. Если дано $C = \{a, b, \ldots\}$ и $F = \{\varrho, \sigma, \ldots\}$ — частично упорядоченные группы, проблема состоит в том, чтобы найти все те частично упорядоченные группы G, в которых C является выпуклым нормальным делителем и факторгруппа G/C (в которой может быть введено естественное упорядочение) изоморфна с F при сохранении порядка. Проблема может быть сведена на системы $C - F$ - расширений $\{f_{\varrho\sigma}, \theta_\varrho, P_{\varrho\sigma}\}$ где $\{f_{\varrho\sigma}\}$ система факторов Шрейера, θ_ϱ автоморфизм группы C которые сохраняют порядок и $P_{\varrho\sigma}$ — подмножества группы C, которые в группе G определяют порядок. Автор доказывает, что если система расширений $C - F$ удовлетворяет условиям (3) — (11), тогда она определяет расширение G на основе $(\alpha) = (\gamma)$. Наконец разбирается лексикографическое и прямое расширение, а также расширение структуры.

ÜBER NIVEAUKURVEN UND FLÄCHEN VON LÖSUNGSFUNKTIONEN PARTIELLER DIFFERENTIALGLEICHUNGEN

Von
J. ACZÉL (Miskolc)
(*Vorgelegt von* G. HAJÓS)

1. Zielsetzung. Im folgenden wird die unter anderen mit der Potentialtheorie und der Nomographie zusammenhängende Frage erörtert, welcher partiellen Differentialgleichung jene Funktionen genügen müssen, deren Niveaukurven- bzw. Niveauflächensysteme mit denen der Lösungsfunktionen einer gegebenen partiellen Differentialgleichung gemeinsam sind. Dieses Problem wird für einige lineare partielle Differentialgleichungen und Systeme erster Ordnung, und homogene Gleichungen und Systeme zweiter Ordnung gelöst.

2/a. Beispiel aus der Potentialtheorie. Eine bekannte Aufgabe in der Potentialtheorie ist die folgende.[1] (Wir beschränken uns der Kürze halber auf Funktionen zweier Veränderlichen.)

Welche sind die notwendigen und hinreichenden Bedingungen dafür, daß eine Kurvenfamilie $F(x, y) = C$, wo $F(x, y)$ wenigstens zweimal derivierbar ist,[2] das Niveaukurvensystem (System äquipotentieller Kurven) einer Potentialfunktion, d. h. einer Lösungsfunktion der Laplace'schen Differentialgleichung

$$U_{xx} + U_{yy} = 0 \tag{1}$$

sei; und wie kann jene Potentialfunktion $U(x, y)$ bestimmt werden, deren äquipotentielle Kurven durch $F(x, y) = C$ angegeben sind? — Die Antwort ist einfach:

Die Forderung, daß $F(x, y) = C$ alle Niveaukurven von $U(x, y)$ und nur diese angebe, d. h. daß das Niveaukurvensystem von $U(x, y)$ mit dem von

[1] Siehe z. B.: O. D. KELLOGG, Foundations of Potential Theory, New York (1929), p. 195.

[2] Auch weiterhin werden wir die vorkommenden Funktionen dem Bedürfnis gemäß als einmal, zweimal oder dreimal derivierbar voraussetzen und, wo es nötig ist, wird eine erste Derivierte als nirgends verschwindend vorausgesetzt.

$F(x, y)$ zusammenfalle, ist offensichtlich mit $U(x, y) = s\,[F(x, y)]$ äquivalent, da ja sonst, wenn U außer F noch z. B. von x abhinge, so könnte U bei $F(x, y) = C$ nicht konstant bleiben. — Wegen (1) ist also

$$0 = U_{xx} + U_{yy} = [s'(F)\,F_x]_x + [s'(F)\,F_y]_y = s'(F)\,(F_{xx} + F_{yy}) + \\ + s''(F)\,(F_x^2 + F_y^2),$$

$$\frac{F_{xx} + F_{yy}}{F_x^2 + F_y^2} = -\frac{s''(F)}{s'(F)} = \varphi(F)\,, \tag{2}$$

$$\varphi(t) = -\frac{s''(t)}{s'(t)}, \qquad s(t) = \int e^{-\int \varphi(t)\,dt}\,dt = \int_c^t e^{-\int_k^t \frac{F_{xx} + F_{yy}}{F_x^2 + F_y^2}\,dF}\,dt\,,$$

$$U(x, y) = s\,[F(x, y)] = \int_c^{F(x,y)} e^{-\int_k^t \varphi(t)\,dt}\,dt = \int_c^{F(x,y)} e^{-\int_k^t \frac{F_{xx} + F_{yy}}{F_x^2 + F_y^2}\,dF}\,dt, \tag{3}$$

und dieses U ist auch fürwahr eine Potentialfunktion, da wegen (2)

$$U_{xx} + U_{yy} = \left[e^{-\int_k^F \varphi(t)\,dt}\,F_x\right]_x + \left[e^{-\int_k^F \varphi(t)\,dt}\,F_y\right]_y =$$

$$= e^{-\int_k^F \varphi(t)\,dt}\,[F_{xx} + F_{yy} - \varphi(F)\,(F_x^2 + F_y^2)] = 0$$

ist, also erwies sich die notwendige Bedingung (2) auch als hinreichend.

Die notwendige und hinreichende Bedingung (2) bedeutet, daß $\dfrac{F_{xx} + F_{yy}}{F_x^2 + F_y^2}$ eine Funktion von $F(x, y)$ allein sein soll. Ist F dreimal derivierbar, so kann dies auch in die Form einer partiellen Differentialgleichung dritter Ordnung umgestaltet werden:

$$\begin{vmatrix} \dfrac{\partial}{\partial x}\left(\dfrac{F_{xx} + F_{yy}}{F_x^2 + F_y^2}\right) & \dfrac{\partial}{\partial y}\left(\dfrac{F_{xx} + F_{yy}}{F_x^2 + F_y^2}\right) \\ F_x & F_y \end{vmatrix} = 0.$$

(3) ergibt die gesuchte Darstellung von U. — Lösung und Beweis erfolgt in gleicher Weise bei Potentialfunktionen von drei oder mehr Veränderlichen. (Siehe 3/c.).

2/b. Beispiel aus der Nomographie. Eine wichtige einfache Klasse von Nomogrammen ist jene, wo alle drei Träger gerade Linien sind.[3] Damit eine

[3] Siehe z. B. M. d'Ocagne: Traité de Nomographie, Paris (1921), Ch. IV, §§ 57, 67, 71, 74; vgl. auch J. Aczél.[4]

Funktion $F(x, y)$ durch ein Nomogramm dieser Art darstellbar sei, ist es notwendig und hinreichend, daß es drei Funktionen $f(x)$, $g(y)$, $s(t)$ gebe derart, daß

$$s[F(x, y)] = f(x) + g(y) \tag{4}$$

gelte.[3]

Es stellt sich die Frage nach einer notwendigen und hinreichenden Differentialgleichungsbedingung für $F(x, y)$, die das Bestehen von (4) sichert, falls $F(x, y)$ genügend oft derivierbar ist.[4, 5]

Nach dem Vorbild der vorhergehenden Nummer können wir derart folgern: Die Funktion $U(x, y) = f(x) + g(y)$ ist bekanntlich durch die Gleichung $U_{xy} = 0$ vollständig charakterisiert. Da aber $U(x, y) = s[F(x, y)]$ ist, haben wir

$$0 = U_{xy}(x, y) = [s(F)]_{xy} = [s'(F) F_x]_y = s'(F) F_{xy} + s''(F) F_x F_y,$$

$$\frac{F_{xy}}{F_x F_y} = -\frac{s''(F)}{s'(F)} = \varphi(F), \tag{5}$$

$$\varphi(t) = -\frac{s''(t)}{s'(t)}, \qquad s(t) = \int e^{-\int \varphi(t)\, dt}\, dt = \int_c^t e^{-\int_k^t \frac{F_{xy}}{F_x F_y} dF} dt\,. \tag{6}$$

Das Gelten der Gleichung (5) ist also notwendig und, wie man leicht einsieht (vgl. 3/a), auch hinreichend zum Bestehen von (4). — (5) bedeutet wieder, daß $\dfrac{F_{xy}}{F_x F_y}$ Funktion von $F(x, y)$ allein ist, und das ist abermals äquivalent mit der partiellen Differentialgleichung dritter Ordnung

$$\begin{vmatrix} \dfrac{\partial}{\partial x}\left(\dfrac{F_{xy}}{F_x F_y}\right) & \dfrac{\partial}{\partial y}\left(\dfrac{F_{xy}}{F_x F_y}\right) \\ F_x & F_y \end{vmatrix} = 0,$$

$$F_x^2 (F_{xy} F_{yy} - F_{xyy} F_y) = F_y^2 (F_{xx} F_{xy} - F_x F_{xxy}). \tag{7}$$

[4] Für die resultierende partielle Differentialgleichung dritter Ordnung (7) siehe P. Saint-Robert: De la résolution de certaines équations à trois variables par le moyen d'une régle glissante, *Memorie d. R. Accad. di Scienze di Torino*, (2) **25** (1871), pp. 53—72; für die Bedingung (5) und für die Bestimmung (6) von $s(t)$ sowie auch von $f(x)$ und $g(y)$ siehe J. Aczél: Zur Charakterisierung nomographisch einfach darstellbarer Funktionen durch Differential- und Funktionalgleichungen, *Acta Scientiarum Math. Szeged*, **12A** (1950), pp. 73—80.

[5] Vgl. auch L. Fantappie: Sulla struttura delle funzioni di più variabili. *Rendiconti di Mat. Roma*, (5) **2** (1941), pp. 61—70 und S. Martis-Biddau: Sulla caratterizzazione di alcune classi di funzioni, *Collectanea Math. Barcelona*, **1** (1948), pp. 67—81, die wahrscheinlich den Zusammenhang der Frage mit der Nomographie und das Resultat (7) von P. Saint-Robert[4] (das sie wiederentdecken) nicht kennen. Da anderseits die Formeln (5), (6) nicht abgeleitet werden, kann dort nur eine indirekte Bestimmung von $s(t)$ angegeben werden.

3/a. Homogene Gleichungen zweiter Ordnung. Der in den vorhergehenden Beispielen benützte Gedankengang läßt sich in folgender Weise verallgemeinern:

Man sucht Bedingungen, die notwendig und hinreichend für das Zusammenfallen des Niveaukurvensystems der wenigstens zweimal derivierbaren Funktion $F(x, y)$ mit dem der Lösungsfunktion von der homogenen partiellen Differentialgleichung

$$(8)\quad a_1(x, y)\, U_x + a_2(x, y)\, U_y + a_{11}(x, y)\, U_{xx} + a_{12}(x, y)\, U_{xy} + a_{22}(x, y)\, U_{yy} = 0$$

sind.

Da unsere Forderung auch hier mit $U(x, y) = s\,[F(x, y)]$ gleichbedeutend ist, ergibt (8)

$$s'(F)\,[a_1 F_x + a_2 F_y + a_{11} F_{xx} + a_{12} F_{xy} + a_{22} F_{yy}] +$$
$$+ s''(F)\,[a_{11} F_x^2 + a_{12} F_x F_y + a_{22} F_y^2] = 0,$$

$$(9)\qquad \frac{a_1 F_x + a_2 F_y + a_{11} F_{xx} + a_{12} F_{xy} + a_{22} F_{yy}}{a_{11} F_x^2 + a_{12} F_x F_y + a_{22} F_y^2} = -\frac{s''(F)}{s'(F)} = \varphi(F),$$

also soll die linke Seite wieder eine Funktion von F allein sein.

Diese Bedingung (9), die für dreimal derivierbares F in der schon erkannten Weise ebenfalls in eine partielle Differentialgleichung dritter Ordnung umgestaltet werden kann, ist aber nicht nur notwendig, sondern auch hinreichend. Dazu brauchen wir nur zu beweisen, daß jeder Funktion F, die der Bedingung (9) Genüge leistet, eine Funktion $U(x, y) = s\,[F(x, y)]$ zugeordnet werden kann, die (mit F gemeinsames Niveaukurvensystem hat und) der Gleichung (8) genügt. Da aus (9)

$$s(t) = \int e^{-\int \varphi(t)\,dt}\, dt$$

ist, muß

$$(10)\quad U(x,y) = \int\limits_c^{F(x,y)} e^{-\int\limits_k^t \varphi(t)\,dt}\, dt = \int\limits_c^F e^{-\int\limits_k^t \frac{a_1 F_x + a_2 F_y + a_{11} F_{xx} + c_{12} F_{xy} + a_{22} F_{yy}}{a_{11} F_x^2 + a_{12} F_x F_y + a_{22} F_y^2}\, dF}\, dt$$

die gesuchte Funktion sein. Und wegen (9) haben wir auch tatsächlich

$$a_1 U_x + a_2 U_y + a_{11} U_{xx} + a_{12} U_{xy} + a_{22} U_{yy} =$$
$$= e^{-\int\limits_k^F \varphi(t)\,dt}\, [a_1 F_x + a_2 F_y + a_{11} F_{xx} + a_{12} F_{xy} + a_{22} F_{yy} -$$
$$- \varphi(F)\,(a_{11} F_x^2 + a_{12} F_x F_y + a_{22} F_y^2)] = 0, \qquad \text{w. z. b. w.}$$

Also haben wir den

SATZ 1. *Die Bedingung (9) oder die mit ihr äquivalente Differentialgleichungsbedingung dritter Ordnung ist notwendig und hinreichend zum Übereinstimmen des Niveaukurvensystems von $F(x, y)$ mit dem von $U(x, y)$, der Lösungsfunktion der Gleichung (8); die Berechnung von $U(x, y)$ aus $F(x, y)$ wird durch die Formel (10) angegeben.*

Ist insbesondere $a_1(x, y) \equiv a_2(x, y) \equiv 0$, so erhält auch (9) die symmetrische Gestalt

$$\frac{a_{11} F_{xx} + a_{12} F_{xy} + a_{22} F_{yy}}{a_{11} F_x^2 + a_{12} F_x F_y + a_{22} F_y^2} = \varphi(F).$$

3/b. Gleichungen erster Ordnung. Ist U eine Lösung der Gleichung

$$a(x, y) + a_1(x, y)\, U_x + a_2(x, y)\, U_y = 0, \tag{11}$$

bzw. der Gleichung

$$a(x, y)\, U + a_1(x, y)\, U_x + a_2(x, y)\, U_y = 0, \tag{12}$$

so erhält man in einer zur bisher beschriebenen ganz analogen Weise den

SATZ 2. *Für das Übereinstimmen des Niveaukurvensystems von $F(x, y)$ mit dem von $U(x, y)$, der Lösungsfunktion der Gleichung (11) bzw. (12), ist die Bedingung*

$$\frac{a(x, y)}{a_1(x, y)\, F_x + a_2(x, y)\, F_y} = \varphi(F)$$

oder die äquivalente Differentialgleichungsbedingung zweiter Ordnung notwendig und hinreichend; und die Formel

$$U(x, y) = -\int_c^{F(x,y)} \varphi(t)\, dt = -\int_c^{F(x,y)} \frac{a}{a_1 F_x + a_2 F_y}\, dF \tag{13}$$

bzw.

$$U(x, y) = e^{-\int_c^{F(x,y)} \varphi(t)\, dt} = e^{-\int_c^{F(x,y)} \frac{a}{a_1 F_x + a_2 F_y} dF} \tag{14}$$

liefert die Berechnung von $U(x, y)$ aus $F(x, y)$.

Übrigens sind (11) und (12) [sowie (13) und (14)] äquivalent, wie die Substitution $\log U(x, y) = V(x, y)$ in (12) sofort zeigt.

3/c. Differentialgleichungssysteme und Differentialgleichungen mit mehr als zwei Veränderlichen. Gleichungssysteme und Gleichungen mit drei oder mehr Veränderlichen, deren Gestalt der der Gleichungen (8), (11), (12) ähnlich ist, lassen sich in der gleichen Weise untersuchen.

Zum Beispiel ist die notwendige und hinreichende Bedingung dafür, daß eine Flächenfamilie $F(x, y, z) = C$ das System äquipotentieller Flächen (das Niveauflächensystem) einer Potentialfunktion $U(x, y, z)$, also einer

Lösung der Laplace'schen Gleichung

$$U_{xx} + U_{yy} + U_{zz} = 0$$

sei:

$$\frac{F_{xx} + F_{yy} + F_{zz}}{F_x^2 + F_y^2 + F_z^2} = \varphi(F) \quad \text{und es ist}^1 \quad U(x, y, z) = \int\limits_c^{F(x,y,z)} e^{-\int\limits_k^t \varphi(t)\, dt}\, dt.$$

Betrachtet man andererseits die Funktionen $F(x, y, z)$, zu denen es vier Funktionen $f(x)$, $g(y)$, $h(z)$, $s(t)$ gibt, derart daß

$$s\,[F(x, y, z)] = f(x) + g(y) + h(z) \tag{15}$$

bestehe, — auch diese Funktionen sind in der Nomographie von Bedeutung —, so erhält man, da die Funktion $U(x, y, z) = s\,[F(x, y, z)]$ durch das Differentialgleichungssystem $U_{xy} = U_{yz} = U_{zx} = 0$ vollständig charakterisiert wird, als notwendige und hinreichende Bedingung für das Bestehen von (15):

$$\frac{F_{xy}}{F_x F_y} = \frac{F_{yz}}{F_y F_z} = \frac{F_{zx}}{F_z F_x} = -\frac{s''(F)}{s'(F)} = \varphi(F)$$

und es ist [5]

$$s(t) = \int e^{-\int \varphi(t)\, dt}\, dt, \qquad U(x, y, z) = \int\limits_c^{F(x,y,z)} e^{-\int\limits_k^t \varphi(t)\, dt}\, dt.$$

4. Probleme. Es stellt sich die Frage, welcher Differentialgleichungsbedingung eine Funktion F genügen muß, damit sie mit einer Funktion U gemeinsames Niveaukurven- bzw. Niveauflächensystem besitze, die Lösungsfunktion einer partiellen Differentialgleichung höherer Ordnung oder anderer Gestalt ist, als die Gleichungen (8), (11), (12).

Ich glaube, daß die Lösung dieses allgemeinen Problems kaum uninteressant oder unbedeutend sein könnte.

Wie man gleich sieht, versagt unsere Methode bei dieser Aufgabe deshalb, weil hier die Einsetzung von $U = s(F)$ in die Differentialgleichnung von U eine Gleichung gibt, deren linke Seite in mehr als zwei Gliederkomplexe zerfällt, die also nicht auf die Gestalt gebracht werden kann, wo ein Ausdruck der Derivierten von F sowie der Koeffizienten eine Funktion von F allein ist.

Durch Differentialgleichungs*systeme* ist allerdings F mit unserer Methode charakterisierbar, auch wenn die Differentialgleichung, deren Lösung U ist, zu einer weit breiteren Klasse von Differentialgleichungen gehört, in der unter anderen alle linearen Differentialgleichungen, deren Koeffizienten von U unabhängig sind, enthalten sind.

Z. B. bekommt man aus $U(x, y) = s\,[F(x, y)]$,

$$a_1(x, y)\, U_x + a_2(x, y)\, U_y + a_{11}(x, y)\, U_{xx} + a_{12}(x, y)\, U_{xy} + a_{22}(x, y)\, U_{yy} = a(x, y) \tag{16}$$

die Gleichung

$$s''(F)\,(a_{11}\,F_x^2 + a_{12}\,F_x\,F_y + a_{22}\,F_y^2) +$$
$$+ s'(F)\,(a_1\,F_x + a_2\,F_y + a_{11}\,F_{xx} + a_{12}\,F_{xy} + a_{22}\,F_{yy}) = a$$

oder mit den Bezeichnungen

$$(17)\qquad u = \frac{a}{a_{11}\,F_x^2 + a_{12}\,F_x\,F_y + a_{22}\,F_y^2}, \qquad v = \frac{a_1\,F_x + a_2\,F_y + a_{11}\,F_{xx} + a_{12}\,F_{xy} + a_{22}\,F_{yy}}{a_{11}\,F_x^2 + a_{12}\,F_x\,F_y + a_{22}\,F_y^2}.$$

$$(18)\qquad s''(F) + s'(F)\,v = u.$$

Derivieren wir dies nach x bzw. y

$$s'''(F)\,F_x + s''(F)\,v\,F_x + s'(F)\,v_x = u_x \qquad F_y$$
$$s'''(F)\,F_y + s''(F)\,v\,F_y + s'(F)\,v_y = u_y \qquad (-\,F_x)$$

$$s'(F) = \begin{vmatrix} u_x & u_y \\ F_x & F_y \end{vmatrix} \Bigg/ \begin{vmatrix} v_x & v_y \\ F_x & F_y \end{vmatrix},$$

d. h. mit der Bezeichnung

$$(19)\qquad w = \begin{vmatrix} u_x & u_y \\ F_x & F_y \end{vmatrix} \Bigg/ \begin{vmatrix} v_x & v_y \\ F_x & F_y \end{vmatrix}$$

ist $\begin{vmatrix} w_x & w_y \\ F_x & F_y \end{vmatrix} = 0$. Dies und (18) ergeben das Gleichungssystem

$$(20)\qquad \frac{w_x}{F_x} = \frac{w_y}{F_y} = u - v\,w.$$

Alle unsere Betrachtungen sind umkehrbar und *so ist das Differentialgleichungssystem (20), wo u, v, w die Bedeutung (17), (19) haben, notwendig und hinreichend zum Bestehen von $U = s(F)$, falls U der Differentialgleichung (16) genügt.*

Ähnlich, bzw. durch wiederholtes Anwenden desselben Verfahrens lassen sich alle Differentialgleichungen der Gestalt

$$\Phi(x, y;\, U, U_x, U_y, U_{xx}, U_{xy}, U_{yy}, \ldots) = 0$$

behandeln, wo Φ ein Polynom von U und seinen Derivierten ist, deren Koeffizienten von x und y abhängen. — Allerdings geht dabei auch die Einfachheit des Ergebnisses verloren.

(Eingegangen am 4. Mai 1950.)

ЛИНИИ И ПОВЕРХНОСТИ УРОВНЯ РЕШЕНИЙ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ В ЧАСТНЫХ ПРОИЗВОДНЫХ

Я. АЦЕЛ (Мишкольц)

(Резюме)

Работа поднимает проблему о том, какому дифференциальному уравнению (или же системе дифференциальных уравнений) должна удовлетворять та функция, у которой система линий поверхностей уровня совпадает с системой линий уровня решений даннного дифференциального уравнения. Нахождением искомого дифференциального уравнения, проблема разрешается в следующих случаях: а) В случае, если первоначальное дифференциальное уравнение 1-го порядка и его левая сторона относительно функции и первой производной однородно линейна; б) если оно 1-го порядка и его левая сторона относительно производных функций однородно линейно, а сама функция явно не фигурирует, и в) если оно 2-го порядка и его левая сторона относительно первой и второй производных функции однородно линейна, а сама функция явно не фигурирует. Проблема разрешается также в более общих случаях. Вопрос связан с известными задачами теории потенциала и номографий.

ON FACTORISATION OF GRAPHS

By
T. GALLAI (Budapest)
(*Presented by* P. TURÁN)

§ 1.

By a *graph* we mean in this paper a set of *vertices* (or *points*) and *edges*, where each edge joins two different vertices called the *endpoints* or *extremities* of the edge[1]. A graph is finite if the set of its points and edges is finite. By a *subgraph* is meant a graph which is a subset of the original graph. An edge will be called *adjoint* to a subgraph if one of its endpoints belongs to the subgraph and the other does not; the former endpoint is said to be the *internal* vertex, the latter one the *external* vertex of the adjoint edge. The number of edges adjoint to a subgraph will be called the *adjoint number* of the subgraph. The number of edges adjoint to a single point is termed the *degree* of the point. If every point has a finite degree, the graph is said to be of finite degree or locally finite. If the degree of all points of the graph is the same, then we say the graph is *regular* and its *degree* is the common degree of its points. A regular subgraph containing each point of the original graph is said to be a *factor* of the graph. In particular, if the graph has a factor of degree 1, then each point of the graph belongs to one and only one edge of the factor; the edges of a factor of degree 2 are „circles" having no vertex in common and containing together every vertex of the original graph[2].

A regular graph having a factor has another factor too, called the *complementary factor*, consisting of the edges not belonging to the factor-graph and of all vertices. The original graph is called the *product* of these two factors. The sum of the degrees of these two factors is clearly the degree of the whole graph. A regular graph having no factors is said to be *primitive*.

The first results of considerable importance concerning the factorisation of graphs are due to PETERSEN. He has proved the following theorems on finite regular graphs[3]:

[1] These concepts will be used here only in the combinatorial sense.
[2] Đ. KÖNIG, Theorie der endlichen und unendlichen Graphen, Akad. Verl., Leipzig (1936), pp. 155—159.
[3] J. PETERSEN, Die Theorie der regulären Graphs, *Acta Math.*, 15 (1891), pp. 193-220.

Every regular graph of even degree has a factor of degree 2. Consequently, a regular graph of even degree can be resolved into the product of factors of degree 2.

There exist primitive graphs of any given odd degree.

If the adjoint number of each proper subgraph of a graph of degree 3 exceeds 1 (i. e. the graph has no „bridge"), then the graph has necessarily a factor of degree 2 (and hence one of degree 1).

The conjecture of PETERSEN concerning factors of degree 2 of graphs of odd degrees greater than 3, as well as an extension thereof to factors of higher degree than 2 were proved by BAEBLER[4]. His theorem is as follows:

A finite graph of degree $2m + 1$ has a factor of degree 2ϱ, if it has no proper subgraph whose adjoint number is less than 2ϱ.

Since in this case, the complementary factor of a factor of even degree is of odd degree, we have also found a condition for the existence of factors of odd degrees. In particular, in case $\varrho = m$ BAEBLER's theorem yields a condition for the existence of a factor of degree 1.

Recently TUTTE[5] has given a necessary and sufficient condition for the existence of a factor of degree 1 of a (not necessarily regular) finite graph. A new proof of this theorem may be found in § 5 of this paper.

KÖNIG[6] has shown that a special type of graphs, the so-called regular graphs of *even circuits*, can be resolved into the product of factors of degree 1. Starting from a theorem due to HALL[7], R. RADO has given a necessary and sufficient condition for the existence of factors of degree 1 in even (not necessarily regular) graphs[8] (see § 5 of this paper).

The cited theorem of KÖNIG has been generalized to infinite even graphs by KÖNIG and VALKO[9], PETERSEN's theorem on even regular graphs by KÖNIG[10], the HALL—RADO theorem by R. RADO[8] and the TUTTE theorem by TUTTE himself[11].

PETERSEN, KÖNIG and BAEBLER obtained their results by using the method of alternating paths. TUTTE used the elegant method of skew-symmetric determinants.

[4] F. BAEBLER, Über die Zerlegung regulärer Streckenkomplexe ungerader Ordnung, *Comment. Math. Helvetici*, **10** (1938), pp. 275—287.

[5] W. T. TUTTE, The factorization of linear graphs, *Journ. London Math. Soc.*, **22** (1947), pp. 107—111.

[6] Loc. cit.[2], pp. 170—175. A graph is said to be of *even circuit*, or, simply *even*, if it has no circle with an odd number of edges.

[7] P. HALL, On representations of subsets, *Journ. London Math. Soc.*, **10** (1934), pp. 26—30.

[8] R. RADO, Factorization of even graphs, *Quart. Journ. Math.* (Oxford series), **20** (1949), pp. 95—104.

[9] D. KÖNIG und S. VALKÓ, Über mehrdeutige Abbildungen von Mengen, *Math. Annalen*, **95** (1926), pp. 135—138.

[10] D. KÖNIG, loc. cit.[2], pp. 203—205.

[11] W. T. TUTTE, The factorisation of locally finite graphs, *Canadian Journ. Math.*, **2** (1950), pp. 44—49.

In the present paper we will show, by further developping the methods of PETERSEN and BAEBLER, how the cited theorems may be derived from a common source. Even for the existence of factors we shall give a condition which is independent of whether the degree of the graph or of the factor is even or odd. Accordingly we shall get in case of graphs of even degree a condition for the existence of factors of odd degrees higher than 1. This problem was not discussed by the aforesaid papers. Our results on adjoint numbers are sharper than the previous ones.

At first all propositions will be proved for the case of finite graphs (in §§ 4—7 only finite graphs are considered), and then in § 8 we extend the theorems simultaneously to infinite graphs of finite degree.

With the present method I have even found a necessary and sufficient condition for the existence of a factor of degree 2; this will be discussed on another occasion.

§ 2.

Our aims necessitate some subdivision of the edges of the graph into classes. Let us therefore consider a graph G, whose edges are arbitrarily divided into two classes. The edges belonging to the first class will be called α-coloured edges, briefly *α-edges*, while those belonging to the other class will be called *β-edges*. The letters u, v, w, x, y, z, as well as $\bar{u}, \bar{v}, \bar{w}, \bar{x}, \bar{y}, \bar{z}$, will serve to denote one of the colours α, β such that

$$x = \alpha \quad \text{implies} \quad \bar{x} = \beta$$
$$\text{and} \quad x = \beta \quad \text{implies} \quad \bar{x} = \alpha.$$

By a *path* we understand a sequence

$$A_0\, e_1\, A_1\, e_2\, A_2 \ldots A_{n-1}\, e_n\, A_n$$

of the vertices $A_0, A_1, \ldots, A_n$ and the *different* edges $e_1, e_2, \ldots, e_n$, where the vertices A_{i-1} and A_i are the extremities of the edge e_i $(i = 1, 2, \ldots, n)$. If in the sequence $e_1, e_2, \ldots, e_n$ the α and β-edges follow in an alternating order, we say the path is an *alternating path*. Since no other path will occur in this paper, we shall say simply path instead of alternating path.

This definition lacks of meaning if $n = 1$. However, we shall regard also the sequence $A_0\, e_1\, A_1$, containing one edge only, as an alternating path, provided the A_0 and A_1 are extremities of the edge e_1; moreover we regard any single vertex of the graph in itself as an (alternating) *path without edge*.

If e_1 is a u-edge and e_n is a v-edge, then, if there is no risk of ambiguity, we shall briefly denote this path in the following manner:

$$A_0\, e_1\, A_1 \ldots A_{n-1}\, e_n\, A_n = (A_0^u - {}^v A_n)$$

and we shall say that A_n *is accessible from* A_0 by a path beginning with a u-edge

and ending in a v-edge, briefly *by a uv-path*. For a path consisting of a single u-edge e_1 we write

$$e_1 = A_0\, e_1\, A_1 = (A_0^u - {}^u A_1).$$

For convenience, we shall regard any vertex as a point that is accessible from itself by a $u\bar{u}$-path ($u = \alpha, \beta$). In other words, we postulate for each point A the existence of the paths without edge $(A^\alpha - {}^\beta A)$ and $(A^\beta - {}^\alpha A)$.

If all vertices and edges of the path $(A_0^u - {}^v A_n)$ belong to a subgraph G' of G, we say that the path is lying in G' and A_n is accessible from A_0 within G' by a uv-path.

If the two paths $(A^u - {}^x B)$ and $(B^{\bar{x}} - {}^v C)$ have no edge in common, then they can be united into one path:

$$(A^u - {}^x B) + (B^{\bar{x}} - {}^v C) = (A^u - {}^v C).$$

Conversely, every path $(A^u - {}^v C)$ may be decomposed in this manner by means of any of its points B. (In case $B = A$ we put $x = \bar{u}$, and if $B = C$, then $x = v$.)

We shall prove the following lemma.

LEMMA. *If some vertex P of the path $(B^x - {}^y C)$ is accessible from the vertex A by a path $(A^u - {}^v P)$, then from A either B is accessible by a ux-path, or C by a uy-path. The new path consists of certain edges of the two old paths.*

Let T denote the first point lying on the path $(A^u - {}^v P)$ which is also a point of the path $(B^x - {}^y C)$. Then we decompose these paths:

$$(B^x - {}^y C) = (B^x - {}^z T) + (T^{\bar{z}} - {}^y C)$$

and

$$(A^u - {}^v P) = (A^u - {}^w T) + (T^{\bar{w}} - {}^v P).$$

The path $(A^u - {}^w T)$ has no edge in common with the path $(B^x - {}^y C)$. Therefore, if $z = \bar{w}$, then

$$(A^u - {}^w T) + (T^z - {}^x B) = (A^u - {}^x B),$$

or, if $z = w$, then

$$(A^u - {}^w T) + (T^z - {}^y C) = (A^u - {}^y C)$$

constitutes a path satisfying the postulates of the lemma.

§ 3.

Let A denote an arbitrary but fixed point of the graph G. Let us consider the vertices of G from the point of view of accessibility from this fixed point A by paths beginning with a β-edge.[12] From now on, if no mention is made

[12] A classification of certain points of a graph from a similar point of view may be found in J. PETERSEN, loc. cit.[3], pp. 211–212 and in D. KÖNIG, loc. cit.[2], pp. 232–233.

of the starting point of a path and the colour of its starting edge, we shall always understand a path issuing from A and beginning with a β-edge. In case a path $(A^\beta - {}^xP)$ exists, we shall call P briefly an *accessible point*. If no such path leading up to P can be found, then P is said to be an *inaccessible point*.

The accessible points can be divided into three classes. Points accessible by a path ending in an α-edge, but not by a path ending in a β-edge, are called *α-points*. Similarly are defined the *β-points*. Points accessible by paths of both kinds are called *γ-points*.

As we have required the existence of the path $(A^\beta - {}^\alpha A)$ without edge we conclude that

(3,1) *A is either an α-point or a γ-point.*

Let us consider an arbitrary edge: $PeQ = (P^u - {}^uQ)$. If P is accessible i. e., a path $(A^\beta - {}^xP)$ exists, then our Lemma in § 2 implies that either a path $(A^\beta - {}^uP)$ or a path $(A^\beta - {}^uQ)$ must also exist. Hence it follows that both points, P and Q, are not simultaneously $\bar{u}$-points, that is, both extremities of a u-edge are not $\bar{u}$-points. In detail this means:

(3,2) *two α-points are not connected by a β-edge;*

(3,3) *two β-points are not connected by an α-edge.*

If, in addition, Q is inaccessible, then of the two former paths only $(A^\beta - {}^uP)$ exists, therefore P is either a u-point or a γ-point. The latter case is impossible, since this would imply the existence of a path $(A^\beta - {}^{\bar{u}}P)$, further, considering that Q is inaccessible, this path could not contain PeQ; consequently, the paths $(A^\beta - {}^{\bar{u}}P)$ and $(P^u - {}^uQ)$ could be united into one path, which contradicts the inaccessibility of Q:

(3,4) *an inaccessible point is not connected with a γ-point, it can be connected with an α-point only by an α-edge or, with a β-point only by a β-edge.*

Let us now consider the subgraph Γ consisting of the γ-vertices of the graph and of all edges joining these vertices. This graph is either *coherent*[13] or falls into coherent parts or, briefly, *blocks*:[14] $\Gamma_0, \Gamma_1, \Gamma_2, \ldots$ (i. e., subgraphs which have no vertex in common):

$$\Gamma = \Gamma_0 \cup \Gamma_1 \cup \Gamma_2 \cup \ldots$$

(3,5) *The external point of an edge adjoint to a Γ_i must be either an α-point or a β-point,*

since it can be neither a γ-point because of the definition of Γ_i, nor an inaccessible point because of (3,4).

[13] In general, a graph is called coherent, if any two points of it are connected by a „path" of the graph (here the „path" is not necessarily alternating).

[14] In PETERSEN, loc. cit. [3], p. 214, „Systeme zweipfeiliger Linien", in BAEBLER, loc. cit.,[4] p. 280, G-komplexes correspond to these blocks.

If we proceed along a path issuing from A to one of the vertices of the block Γ_i, we have to enter the block through an edge adjoint to this block (not to mention the case when the whole path together with A lies in block Γ_i). If the external point of such an adjoint edge is an α-point, then the edge can be nothing el e but a β-edge; if on the other hand the external point is a β-point, the edge must be an α-edge. In what follows we shall call every edge adjoint to the block Γ_i, whose external point is a u-point and which is a $\bar{u}$-edge, an *entering edge* of the block.

The following theorem is intended to serve as the main tool in later considerations and theorems.

THEOREM I. *There is at most one entering edge adjoint to a block Γ_i. If Γ_i does not contain A, there is one entering edge, in case Γ_i does contain A, there is none.*

Let us consider a path $(A^\beta - {}^xR)$ leading to some point R of Γ_i and let the first point of this path which belongs to Γ_i be denoted by A_i. Then $(A^\beta - {}^xR) = (A^\beta - {}^{\bar{v}}A_i) + (A_i^v - {}^xR)$, where the path $(A^\beta - {}^{\bar{v}}A_i)$ has no edge belonging to Γ_i. We shall examine the accessibility of the points of Γ_i from A_i by paths beginning with a v-edge and lying in Γ_i. (In what follows, when speaking of *accessibility within Γ_i* we shall confine ourselves to the accessibility from A_i by a path beginning with a v-edge and lying in Γ_i.)

Let us first prove two auxiliary theorems.

AUXILIARY THEOREM 1. *If some point R of Γ_i is accessible within Γ_i by a path ending in a y-edge, then it is also accessible within Γ_i by a path ending in a $\bar{y}$-edge.*

Assuming the existence of a path $(A_i^v - {}^yR)$ in Γ_i we shall prove the existence of a path $(A_i^v - {}^{\bar{y}}R)$ lying in Γ_i.

As R is a γ-point, there exists a path $(A^\beta - {}^{\bar{y}}R)$. If all edges of this path belong to Γ_i, then all its points together with A belong to Γ_i; then $(A^\beta - {}^{\bar{y}}R)$ is a desired path, since $A_i = A$ and $v = \beta$.

If the path $(A^\beta - {}^{\bar{y}}R)$ has an edge not belonging to Γ_i, then let $e = MeN = (M^z - {}^zN)$ denote the last of such edges. e is evidently an edge adjoint to Γ_i. M is its external point and N is its internal point. We have

$$(A^\beta - {}^{\bar{y}}R) = (A^\beta - {}^{\bar{z}}M) + (M^z - {}^zN) + (N^{\bar{z}} - {}^{\bar{y}}R),$$

where the path $(N^z - {}^yR)$ lies in the block Γ_i. Owing to our Lemma, from the existence of the paths $(N^z - {}^{\bar{y}}R)$ and $(A_i^v - {}^yR)$ lying in Γ_i we may conclude that there must exist either a path $(A_i^v - {}^{\bar{y}}R)$ lying in Γ_i (in this ca e the auxiliary theorem has already been proved) or a path $(A_i^v - {}^{\bar{z}}N)$ lying in Γ_i. In the latter case the edge e must belong to the path $(A^\beta - {}^{\bar{v}}A_i)$, for otherwise the paths $(A^\beta - {}^{\bar{v}}A_i)$, $(A_i^v - {}^{\bar{z}}N)$ and $(N^z - {}^zM)$ having no

edge in common could be united into one path. This is impossible, because in view of the path $(A^\beta - {}^{z}M)$ the point M would be accessible from A by a path ending in a z-edge as well as by one ending in a $\bar{z}$-edge; in other words, M would be a γ-point contrary to (3,5). This shows that in the second alternative the edge e belongs in fact to the path $(A^\beta - {}^{\bar{v}}A_i)$. But this can happen only if $N = A_i$ and $z = \bar{v}$. Then

$$(N^{\bar{z}} - {}^{\bar{y}}R) = (A_i^v - {}^{\bar{y}}R)$$

is a desired path.

Auxiliary theorem II. *Every point of Γ_i is accessible within Γ_i both by paths ending in α-edges and by those ending in β-edges.*

Owing to auxiliary theorem I all we have to prove is that every point of Γ_i is accessible within Γ_i by some path.

Let us consider all points of Γ_i which are accessible within Γ_i and let us denote the subgraph consisting of these points and of the edges joining these points by $\bar{\Gamma}_i$. Since the path $(A_i^v - {}^{\bar{v}}A_i)$ (without edge) exists, A_i is accessible within Γ_i: therefore $\bar{\Gamma}_i$ is not void. Assuming that there is a point inaccessible within Γ_i, i. e. $\bar{\Gamma}_i \neq \Gamma_i$, from the coherence of Γ_i it follows that there is a point K, inaccessible within Γ_i, connected with a point L accessible within Γ_i. The joining edge $f = KfL = (K^x - {}^xL)$ does not belong to $\bar{\Gamma}_i$. In accordance with auxiliary theorem I L is accessible within Γ_i by paths of any ending edge, therefore a path $(A_i^v - {}^xL)$ in Γ_i must exist. All its vertices and edges are also in $\bar{\Gamma}_i$ and thus the edge f can be united with it in the following manner:

$$(A_i^v - {}^xL) + (L^x - {}^xK) = (A_i^v - {}^xK).$$

This is, however, contradictory to the inaccessibility of K within Γ_i. Consequently, there is no point inaccessible within Γ_i, q. e. d.

Using the auxiliary theorems we can prove theorem I as follows.

Let us assume that the edge $g = (S^u - {}^uT)$ is one of the entering edges of the block Γ_i. If S is the external point of this edge, then S is a $\bar{u}$-point. The internal point T is accessible from A_i, according to auxiliary theorem II, by a path $(A_i^v - {}^{\bar{u}}T)$ in Γ_i. The paths $(A^\beta - {}^{\bar{v}}A_i)$, $(A_i^v - {}^{\bar{u}}T)$ and $(T^u - {}^uS)$ can not be united into one path $(A^\beta - {}^uS)$, for S would be in this case accessible by a path ending in a u-edge, i. e., it would not be a $\bar{u}$-point. The only case when these three paths can not be united into a single path is clearly that when g is an edge of the path $(A^\beta - {}^{\bar{v}}A_i)$. This is inconsistent if $A_i = A$: therefore, if Γ_i contains A, no entering edge exists. If $A_i \neq A$, then the only case when g lies in the path $(A^\beta - {}^{\bar{v}}A_i)$ is when $T = A_i$ and g is the last edge of the path. Hence it follows that g is the only entering edge of Γ_i, q. e. d.

The above discussions imply that

(3,6) *each block Γ_i contains more than one point.*

For auxiliary theorem II ensures the existence of a path $(A_i^\nu - {}^\nu A_i)$ in Γ_i. This is not without edge and therefore it contains more than one point.

§ 4.

Now we come to the problem of examining the condition for the existence of a factor of degree $\sigma \geqq 1$ of a finite graph[15]. We have to find a condition for the case when the edges of the graph G can be coloured by the colours α and β so that in each point the number of the incident α-coloured edges shall be σ.

The edges in any graph can certainly be coloured so that the number of the α-edges incident to each vertex of the graph is at most σ („*admissible colouring*"). We shall henceforth allow such a type of colouring only.

If the number of α-edges incident to one vertex is just σ, we shall call this point *saturated*, and if it is less than σ, we call the point *unsaturated*. In the latter case the number of the missing α-edges will be called the *deficiency* of the vertex.

Let us now consider an admissible colouring for which the sum of the deficiencies taken for all points is minimal. Such a colouring necessarily exists by the finiteness of the graph.

No „*singular path*" may exist in case of such a colouring, i. e., no path $(B^\beta - {}^\beta C)$ where B and C are unsaturated and the sum of its deficiencies exceeds 1 (the latter requirement guarantees that if $B = C$, the deficiency in this point is greater than 1). In fact, changing the colours of the edges in such a path, we diminish the deficiencies with 1 in B and C, and leave them unaltered in the other points. The new colouring is again an admissible one because of the properties of the singular path. Here the deficiency sum would be 2 less than the minimal, which is absurd.

Furthermore let us assume that

(4, H) t h e g r a p h G h a s n o f a c t o r o f d e g r e e σ,

i. e., no colouring exists, where every point is saturated, or otherwise, the minimal sum of deficiencies is greater than 0. We show that assuming this hypothesis we shall obtain results contradictory to some other assumptions and therefore, these assumptions will assure the existence of a factor of degree σ.

We shall transfer the notations and statements of the previous sections to the colouring of a graph G with a minimal sum of deficiencies, so that *A shall always stand for an unsaturated point of the graph.*

[15] Since in §§ 4—7 we confine ourselves exclusively to finite graphs, the term „finite" will in general be suppressed.

What has been said about the singular path implies

(4,1) *every β-point and every γ-point is saturated with the only exception of the point A if A is a γ-point and in this case the deficiency in A is exactly 1.*

Since the factorisation problem is essentially simpler in case $\sigma = 1$, and therefore we can establish more general theorems, first we shall deal with this case separately.

§ 5.

We consider now the case $\sigma = 1$. Since apart from A, to an α-point at least one α-edge must be incident, the α-points are now saturated. To every α-point and by (4,1) even to every β- and γ-point (except A) exactly one α-edge is incident. We can also conclude that

(5,1) *an α-edge issuing from an α-point ends neither in an α-point nor in an inaccessible point.*

Indeed, an edge $(P^{\alpha} - {}^{\alpha}Q)$ is the *only* α-edge incident to both P and Q. If P is an α-point, i. e., there exists a path $(A^{\beta} - {}^{\alpha}P)$, then the last edge in this path is necessarily $(Q^{\alpha} - {}^{\alpha}P)$. Leaving this edge aside, the path $(A^{\beta} - {}^{\beta}Q)$ remains. But this means that Q is neither an α-point nor an inaccessible point, as stated.

(5,2) *The entering edge adjoint to a Γ_i is an α-edge.*

The internal point A_i of an entering β-edge would be — in accordance with auxiliary theorem II — accessible from A_i within Γ_i by an $\alpha\alpha$-path. (In the present case $\bar{v} = \beta$, $v = \alpha$.) Then two α-edges would be incident to A_i, what is impossible because of $\sigma = 1$. Similarly, auxiliary theorem II implies that every point of Γ_i is accessible within Γ_i by a path ending in an α-edge, therefore each point of Γ_i with the sole exception of A_i is an endpoint of some α-edge belonging to Γ_i. Hence it follows on the one hand that

(5,3) *besides the entering α-edge, only β-edges are adjoint to Γ_i,*

and on the other hand that

(5,4) *every Γ_i consists of an odd number of points,*

viz. it consists of the point A_i and of an even number of extremities of α-edges having no vertex in common.

Consequently, the graph will have the following structure:

A β-edge issuing from an α-point ends neither in an α-point (3,2), nor an inaccessible point (3,4), nor a γ-point (5,2); hence its endpoint is a β-point.

An α-edge issuing from an α-point terminates neither in an α-point, nor an inaccessible point (5,1), nor a γ-point (5,3); thus its endpoint is a β-point.

The external point of a β-edge adjoint to a block Γ_i is neither an α-point (5,2), nor an inaccessible point (3,4), nor, by definition of Γ_i, a point belonging to another block, consequently it is a β-point.

The single α-edge adjoint casually to a block Γ_i must start, by (5,2) and by definition of Γ_i, from a β-point.

By (4,1), to each β-point an α-edge is incident which ends neither in a β-point (3,3), nor in an inaccessible point (3,4), i. e., it ends either in an α-point or is the entering edge of some block Γ_i.

These discussions show that, leaving A out of consideration, each α-point is the endpoint of one and only one α-edge issuing from a β-point, and each block (with the exception of the block containing A, if this exists) has one and only one α-edge of the same type as its sole entering edge. Hence it follows that there is a one-to-one correspondence between the set of all α-points and blocks Γ_i (except A resp. the block containing A) and the set of β-points. If we omit now the β-points and every edge issuing from them, then the α-points become isolated points and the blocks Γ_i lose their adjoint edges. The α-points and the blocks Γ_i are coherent parts of the remaining graph and each of them contains an odd number of points. But the above correspondence shows that these parts are, just because of the existence of A, in number one more that the β-points omitted.

Thereby we have given a new proof for the sufficiency of TUTTE's theorem:[5]

If a graph has no factor of degree 1, then there is a subset (void or not) of the points of the graph such that by omitting it together with the edges incident to these points, we arrive at a graph whose coherent parts containing an odd number of points are in number more than the number of points omitted.

The converse of this theorem also holds; the following proof is due to TUTTE.[5] If a graph has a subgraph with stated properties, and a factor of degree 1 existed, then to each part arising by the omission we should have to find an adjoint edge of the original graph, which is also an edge of the factor of degree 1. This follows from the fact the number of the points is odd and therefore the edges of the factor incident to the points of a part can not all be edges of the same part. Clearly, the external endpoint of this adjoint edge must be one of the points omitted. But this is impossible, since the number of the points omitted is less than the number of the parts and two edges of a factor of degree 1 have no point in common.

We also give another formulation of TUTTE's theorem; for this purpose we introduce some new terminologies which we shall use later.

Two subgraphs containing no common points, of a graph are *connected* if there is an edge whose endpoints are in the first and in the second subgraph, respectively. A point is connected with a system of subgraphs, if it is

connected with at least one of them. We shall speak of *independent subgraphs*, or of a set of independent subgraphs, if no two of them have a point in common and no two are connected.

Now TUTTE's theorem may be reformulated as follows.

A necessary and sufficient condition that a finite graph shall have a factor of degree 1, is that the number of points connected with any system of independent subgraphs containing an odd number of points shall not be less than the number of the subgraphs in the system.

Let us now assume that the graph in question is even.

It is a well-known elementary fact[16] that the points of such a graph may be divided into two groups in such a manner that no two points belonging to the same group are connected by an edge. The edges of the graph are therefore edges connecting two points of different groups. Thus if we start from A along a path with a β-initial edge, we arrive alternately at points of the two groups. Therefore the points accessible by a path ending in an α-edge belong to the same group as A, and the β-points to the other group. Hence we conclude that the α-points are independent and the graph fails to have γ-points. (Both assertions hold even if $\sigma > 1$.) This implies that the α-edges issuing from β-points terminate in α-points. As we have seen, the α-points are connected neither with each other, nor with inaccessible points; consequently, these points (together with A) form a system of independent points, which is connected only with β-points whose number is one less than that of this system. Therefore, if an even graph has no factor of degree 1, then it has a system of independent points, connected with points in number less than the number of the points in the system.

On the other hand, it is trivial that the existence of a factor of degree 1 implies the possibility of finding to any choice of independent points, points in number equal to the chosen points such that any of them is connected with some of the chosen points. For, if we consider the edges of the factor of degree 1 which are incident to the chosen points, the other endpoints of these edges have the required property.

This completes a new proof of the HALL—RADO theorem:[7,8]

A necessary and sufficient condition for the existence of a factor of degree 1 of a finite even graph is that with any system of independent points at least as many points shall be connected as the number of the points in the system.

Thus we see that, against TUTTE's theorem, in case of even graphs it is sufficient to consider only systems of independent points instead of independent systems of subgraphs containing an odd number of points.

[16] KÖNIG, loc. cit.[2], pp. 151 and 170.

§ 6.

Let us turn our attention to the case of factors of degree $\sigma > 1$. Our method applies in case $\sigma > 2$ only for regular graphs[17]. Therefore, in the following chapter we assume that the finite graph G is a regular graph of degree $\nu > 1$. At the same time we may assume — without loss of generality — that the graph is coherent. By the regularity, if there is a factor of degree σ, then its complementary factor exists and is of degree $\tau = \nu - \sigma$.

Since the number of points of odd degree of a graph is even, it follows that if at least one of σ and τ is odd, then factors of degree σ resp. τ can exist only in case G contains an even number of points. If $\nu = \sigma + \tau$ is also odd, then this condition is automatically satisfied. However, we assume that

(6.C) *if both σ and τ are odd, then the graph has an even number of points.*

Now we are going to show that, if we proceed again by assuming the failure of factors of degree σ (4.H), then by making use of the regularity and coherence of the graph as well as of condition (6.C) we are led to a contradiction arising against the hypothesis that the adjoint numbers of proper subgraphs exceed certain bounds. The fulfilment of these conditions will therefore ensure the existence of a factor of degree σ (and at the same time one of degree τ).

In order to establish the structure of the graph we shall introduce a few more notations.

We shall denote the block Γ_i containing A (inasmuch as such a block exists) by Γ_A, and the number of α- and β-edges adjoint to Γ_A by λ_α and λ_β respectively (if Γ_A does not exist, we put $\lambda_\alpha = \lambda_\beta = 0$).

We shall write Γ_{α_i} for the blocks to which an α-coloured entering edge belongs, and Γ'_{α_i} for those Γ_{α_i} to which no α-edge other than the entering edge is adjoint. g_α will denote the number of the blocks Γ_{α_i}, and g'_α the number of the blocks Γ'_{α_i}. The marks Γ_{β_i}, Γ'_{β_i}, g_β, g'_β have similar meanings.

Let n_α denote the number of α-points and n_β that of β-points. Let further ε mean 1 if A is an α-point and 0 if A is a γ-point (cf. (3.1)).

We are going to show that

(6.1) *each block Γ'_{α_i} or Γ'_{β_i}, if they exist, contains an odd number of points.*

Let us consider, for instance, the α-edges of a block Γ'_{α_i}, leaving all other α-edges aside. If σ is odd, then of the points of Γ'_{α_i} only the entering point (i. e., the internal endpoint of the entering edge) is of even degree in the α-edges, for all points being γ-points are saturated (4.1). Then the fact that the number of the points of odd degree is even implies (6.1). If σ is even,

[17] With the present method I have succeeded in getting factorisation theorems for general graphs besides $\sigma = 1$ only for the case $\sigma = 2$. I shall discuss these results on another occasion.

then only the entering point might be of odd degree in the α-edges. This is, however, impossible and therefore we have

(6,2) *if σ is even, then no block Γ'_{α_i} exists, i. e. $g'_\alpha = 0$.*

The same inference shows that

(6,3) *if τ is even, no block Γ'_{β_i} exists, i. e. $g'_\beta = 0$.*

The existence of Γ_A, i. e., the fact that A is a γ-point, implies the existence of an edge adjoint to it. For, in the contrary case, in view of (4,1) we could show as before that both σ and τ are odd and Γ_A has an odd number of points. Because of the coherence of G we have then $G = \Gamma_A$ in contradiction to our assumption (6,C). Therefore we may state that if Γ_A exists ($\varepsilon = 0$), then $\lambda_\alpha + \lambda_\beta > 0$ and hence

(6,4) *ε, λ_α and λ_β do not vanish simultaneously.*

Let us denote by ξ the minimum of the adjoint numbers of those proper subgraphs of the graph, which contain an odd number of points. The coherence implies $\xi \geqq 1$.

We shall count in two ways the α-edges incident to α-points[18] and the β-edges incident to β-points and then compare the results.

First method of counting. The number of the α-edges incident to α-points is by no means greater than $n_\alpha \sigma - \varepsilon$.

Second method of counting. The number α-edges issuing from a β-point is exactly σ [cf. (4,1)]. The α-edges leading from β-points to γ-points are the entering edges of the blocks Γ_{α_i}, their number is g_α (Theorem I). The other α-edges issuing from β-points end in α-points [(3,3) and (3,4)]. Hence the number of α-edges connecting α-points with β-points is $n_\beta \sigma - g_\alpha$.

The α-edges adjoint to Γ_A end in α-points, this follows from Theorem I and (3,4). Their number is λ_α.

In addition to the entering edge, at least one more α-edge is adjoint to every Γ_{α_i} which is not a Γ'_{α_i}. These end in α-points [Theorem I and (3,4)]; their number is at least $g_\alpha - g'_\alpha$.

By the definitions of the blocks Γ'_{β_i} and the number ξ, the number of α-edges adjoint to a block Γ'_{β_i} is at least $\xi - 1$ and they end in α-points [Theorem I and (3,4)]. Their number is at least $g'_\beta (\xi - 1)$.

Collating the two results, it is required that

$$(n_\beta \sigma - g_\alpha) + \lambda_\alpha + (g_\alpha - g'_\alpha) + g'_\beta (\xi - 1) \leqq n_\alpha \sigma - \varepsilon,$$

whence

(6,5) $$\varepsilon + \lambda_\alpha + g'_\beta (\xi - 1) - g'_\alpha \leqq (n_\alpha - n_\beta)\, \sigma .$$

[18] An α-edge incident to β-points on both ends is counted twice; the same is true for the β-edges.

Turning our attention to the β-edges, *by the first method of counting* the number of β-edges incident to β-points is $n_\beta \tau$.

By the second method the number of β-edges issuing from an α-point is at least τ. (If A is an α-point, then it is for this point at least $\tau + 1$.) g_β of these are adjoint to some of $\Gamma_{\beta i}$, the rest of them ends in β-points [Theorem I, (3,2) and (3,4)]. Hence the number of the latter ones is at least $n_\alpha \tau + \varepsilon - g_\beta$.

The same inference as above shows that the numbers of β-edges in the corresponding three cases are λ_β, at least $g_\beta - g'_\beta$ and at least $g'_\alpha(\xi - 1)$ respectively.

By comparing the results we get

$$(n_\alpha \tau + \varepsilon - g_\beta) + \lambda_\beta + g'_\alpha(\xi - 1) + (g_\beta - g'_\beta) \leqq n_\beta \tau ,$$

whence

$$\varepsilon + \lambda_\beta + g'_\alpha(\xi - 1) - g'_\beta \leqq (n_\beta - n_\alpha)\tau. \tag{6,6}$$

By eliminating n_α and n_β from (6,5) and (6,6) we get

$$\varepsilon(\sigma + \tau) + \tau\lambda_\alpha + \sigma\lambda_\beta + g'_\alpha[\sigma(\xi - 1) - \tau] + g'_\beta[\tau(\xi - 1) - \sigma] \leqq 0 . \tag{6,7}$$

According to (6,4), since $\sigma > 0$, $\tau > 0$, we obtain

$$g'_\alpha[\sigma(\xi - 1) - \tau] + g'_\beta[\tau(\xi - 1) - \sigma] < 0 . \tag{6,8}$$

We find a contradiction to our hypothesis (4,H) if we stipulate the inequalities

$$\sigma(\xi - 1) - \tau \geqq 0 \qquad \text{and} \qquad \tau(\xi - 1) - \sigma \geqq 0, \tag{6,9}$$

i. e.,

$$\xi \geqq \frac{\nu}{\sigma} \qquad \text{and} \qquad \xi \geqq \frac{\nu}{\tau} .$$

We have thereby proved the following theorem:

THEOREM II. *A coherent regular finite graph of degree $\nu = \sigma + \tau$ can be decomposed into the product of factors of degree σ and of degree τ, if the adjoint numbers of its proper subgraphs containing an odd number of points exceed or equal the greater of the numbers $\frac{\nu}{\sigma}$ and $\frac{\nu}{\tau}$; provided that whenever both σ and τ are odd, the graph has an even number of points.*

(6,9) is equivalent to

$$\frac{\nu}{\xi} \leqq \sigma \leqq \nu - \frac{\nu}{\xi} . \tag{6,10}$$

Theorem II can be reformulated as follows:

COROLLARY 1. *The graph has a factor of degree σ for every number σ satisfying the following condition:*

$$\frac{\nu}{\xi} \leqq \sigma \leqq \nu - \frac{\nu}{\xi}$$

provided that, in case σ is odd and ν is even, the graph has an even number of points.

§ 7.

By considering the evenness resp. oddness of the numbers σ, τ and ν, we can further sharpen the above results.

a) (6,2) and (6,3) show that if *both σ and τ are even*, then

$$g'_\alpha = g'_\beta = 0.$$

This is contradictory to (6,8) and to assumption (4,H). This implies Petersen's theorem[3] on graphs of even degree:

Every finite graph of even degree may be splitted into the product of two factors of arbitrary even degrees. (Of course, the sum of the degrees of the factors must be equal to the degree of the graph.)

b) *If σ is even and τ is odd,* then we can ascertain the vanishing of g'_α only. Then the inequality

$$\tau(\xi - 1) - \sigma \geqq 0$$

resp. the hypothesis

$$\xi \geqq \frac{\nu}{\nu - \sigma} \qquad \text{or} \qquad \sigma \leqq \nu - \frac{\nu}{\xi} \tag{7,1}$$

contradicts (6,8) and so even (4,H). Hence we have

Theorem III. *A finite graph of an odd degree ν has a factor of even degree σ if ξ satisfies condition* (7,1).

Now we prove the following

Lemma. *ξ is even or odd according as ν is even or odd.*

All subgraphs of a graph of even degree have an even adjoint number. For let us consider an arbitrary subgraph and let us complete it by all those edges of the original graph whose both endpoints belong to this subgraph. Then unite all external points of the edges adjoint to this subgraph into one point K. All points different from K of the graph thus arising are of even degree and thus K cannot be the only point of odd degree. Hence the adjoint number of each subgraph is even, in fact. Similar arguments show that a subgraph having an odd number of points, of a graph of odd degree, has an odd adjoint number. This completes the proof of the lemma.

In view of this lemma we see that for graphs of an odd degree $\xi > 1$ implies $\xi \geqq 3$ and from Theorem III we conclude

Theorem IV. *If a finite graph of an odd degree ν has no „bridge" ($\xi > 1$), then it has always a factor of any even degree not exceeding $\frac{2}{3}\nu$, consequently, also a factor of any odd degree not less than $\frac{1}{3}\nu$.*

If $\nu = 3$, we get PETERSEN's theorem quoted in § 1.

Since by $0 < \sigma < \nu$

$$\frac{\nu}{\nu - \sigma} \leqq \sigma + 1,$$

the existence of (7,1) follows from the condition $\xi \geqq \sigma + 1$. On the other hand, if ξ is odd and σ is even, it is enough to suppose $\xi \geqq \sigma$. Accordingly, we have

A finite graph of odd degree has a factor of even degree σ if $\xi \geqq \sigma$.

This is BAEBLER's theorem cited in § 1.

c) *If both σ and τ are odd,* then ν and hence by lemma ξ is even. In this case for instance $\xi > 2$ implies $\xi \geqq 4$ and so, if $\xi > 2$, by (6,10) factors of degree σ can be found for every σ satisfying the inequality $\frac{\nu}{4} \leqq \sigma \leqq \frac{3\,\nu}{4}$. The condition $\xi > 2$ can be replaced by the condition that the graph has no „double bridge", i. e., by omitting any two edges the coherence of the graph will not be destroyed. Consequently, we have proved

THEOREM V. *If a coherent finite graph of even degree with an even number of points has no double bridge, then it has a factor of degree σ for every σ satisfying the inequality*

$$\frac{\nu}{4} \leqq \sigma \leqq \frac{3\,\nu}{4}.$$

We remark that in case of an even ν it follows $\xi \geqq 2$ and hence the inequality

$$\frac{\nu}{\xi} \leqq \frac{\nu}{2} \leqq \nu - \frac{\nu}{\xi}$$

is fulfilled without any further assumption. This means:

A finite graph of an even degree ν with an even number of points always possesses a factor of degree $\frac{\nu}{2}$. (KÖNIG.[19])

As it has been shown in § 5, an even graph contains no γ-points, hence Γ_{α_i} and $\Gamma_{\beta_i'}$ fail to exist. Consequently, $g'_\alpha = g'_\beta = 0$ is always true, leading at once to a contradiction to (4,H) and proving that the graph has a factor of degree σ whatever the value of σ ($\sigma \leqq \nu$). Since A is necessarily an α-point, ε is equal to 1 and so we have not even to assume[20] (6,C) to prove (6,4). We thus get

[19] KÖNIG, loc. cit. [2], p. 160.

[20] We remark that for for regular even graphs, as easily proved, (6, C) is automatically fulfilled.

A regular finite even graph possesses a factor of any degree not exceeding the degree of the graph. (König.[21])

§ 8.

In this section we extend the results of the previous sections to infinite graphs of finite degree. Since a coherent graph of finite degree may have only denumerable points[22], in what follows we may confine ourselves only to graphs containing denumerable points. Evidently, in an infinite graph of finite degree the edges adjoint to any finite subgraph are finite in number.

We are going to prove the following theorem.

Theorem VI. *Suppose that an infinite graph G of finite degree satisfies the conditions for the existence of factors of degree σ of finite graphs stated in §§ 5—7 (with the trivial exception of the condition of containing an even number of points) such that by „systems of points," „subgraphs," „coherent parts" of G we mean only the finite ones. Then G has a factor of degree σ.*

This theorem essentially contains as special cases the theorems of König—Valkó, König, Radó and Tutte cited in §1.

In the preceding discussions the finiteness of the graph was first used in § 4, when we assumed the existence of a minimal deficiency-sum. Hence we concluded that the graph has no singular path and this led us to (4,1) on which all subsequent statements are based. Besides this we used the finiteness of the graph only through the finiteness of the number of accessible points in the graph.

Starting with the assumptions and notations of §3 we shall show, applying a reasoning due to König and Valkó[9], that every infinite graph of finite degree has a colouring such that the number of α-edges incident to any point does exceed σ and there is no (finite) singular path in the sense of §4. We shall call such colourings *suitable*.

Let $P_1, P_2, \ldots$ denote the points of the graph in some order. By G_n $(n = 1, 2, \ldots)$ we shall mean the subgraph consisting of the points $P_1, P_2, \ldots, P_n$ and the edges connecting these points. If $k < l$, G_k is a subgraph of G_l; each colouring of G_l implies a colouring of G_k which will be called a part of the colouring of G_l, and the colouring of G_l is said to be an extension of the colouring of G_k. (We observe that for the sake of uniformity we shall also speak of the colouring of graphs without edge, this having all other colourings as its extension.) It is evident that each finite graph G_n has admissible colourings with a minimal deficiency-sum, i. e. suitable colourings, but only a finite number of them.

[21] König, loc. cit. [2], pp. 170—195.

[22] König, loc. cit. [2], pp. 73—80.

The suitable colourings of the graphs $G_2, G_3, \ldots$ imply suitable colourings of G_1. Therefore, G_1 has a colouring which is part of an infinity of suitable colourings. Let us fix one of such colourings of G_1, and in what follows we consider only the extensions of this fixed colouring. Each of these implies a suitable colouring on G_2, hence G_2 has a suitable colouring which is part of infinitely many suitable colourings, etc. We proceed in this manner to get fixed suitable colourings of the graphs $G_1, G_2, \ldots$ such that each is an extension of its predecessor. Therefore, each edge of G obtains a well-defined colour and this defines a colouring of G which is evidently suitable (since a singular path of G would be that of some G_n).

Assume now that the graph satisfies the conditions of Theorem VI. We shall arrive at a contradiction to the hypothesis that

(8,H) the graph has no factor of degree σ.

Owing to (8,H), in every suitable colouring the graph must have unsaturated points. But unsaturated points may only occur such that the points accessible from any fixed one of them on a path with an initial β-edge are infinite in number. For, if an unsaturated point existed from which only finitely many points were accessible by paths with initial β-edges, then choosing this point for the point A of § 4, (4,1) would be true because of the suitability of the colouring and we should arrive, applying the reasonings of §§ 5—7, at a contradiction to the hypotheses of Theorem VI.

Let us choose an arbitrary suitable colouring and let the unsaturated points be $A_1, A_2, \ldots$ (this is either a finite or an infinite sequence). From each of these points infinitely many points are accessible by a path beginning with a β-edge. Now we show that in this case from each of these points an infinite alternating path starts, having an initial β-edge.

Consider A_j and the infinitely many alternating paths issuing from A_j and beginning with a β-edge. If $l > n$, then a path consisting of l edges contains a subpath of n edges. Because of the finite degree of the graph, the number of paths containing n edges are finite in number. Hence, there is a path with n edges which is part of infinitely many paths, and applying the same inference as above, we conclude that there is an infinite alternating path issuing from A_j with an initial β-edge.

If we interchange the colours of the edges on such a path, we diminish the deficiency of A_j without altering the deficiencies of the other points. We shall prove at the end of this section that the new colouring got by interchanging the colours is again a suitable colouring of the graph. Suppose that we have already verified this statement.

At first, interchange the colours of an infinite path issuing from A_1. If A_1 is yet unsaturated, do the same with another path of the same type etc. until A_1 becomes saturated. Then we apply the same process on $A_2, A_3, \ldots$

If the number of A_i were finite, then we would arrive after a finite number of steps at a colouring, where each A_i, and hence each point of the graph is saturated. This is against $(8,H)$. We show now that, in case of infinitely many A_i, one is able to construct a colouring of the graph where each point is saturated, i. e., Theorem VI is true.

We shall denote the colourings got by successive interchanging of colours as described above by $\varphi_1, \varphi_2, \ldots$ It is clear that to any k we can find an N such that in φ_n $(n > N)$ all points $A_1, \ldots, A_k$ are saturated. Since these infinitely many colourings determine on each subgraph G_i only finitely many different subcolourings, by the König—Valkó method we can successively fix on the subgraphs $G_1, G_2, \ldots$ colourings $\psi_1, \psi_2, \ldots$ such that each ψ_{i+1} is an extension of ψ_i and each of them is a part of infinitely many φ_j. The colourings $\psi_1, \psi_2, \ldots$ determine uniquely a colouring ψ of the edges of G where each point is saturated. For, any point P of the graph, together with the extremities of the edges issuing from it, belongs to some subgraph G_j.[1] Since ψ_j, the colouring of G_j determined by ψ, is part of infinitely many φ_i, it is part of a certain colouring φ_i, at which the points A_k belonging to G_j (finite in number) and hence all the points of G_j are saturated together with P. But every edge issuing from P belongs to G_j, hence P is saturated in ψ_j and so even in ψ.

In order to complete the proof we have to show that if we interchange the colours on an infinite alternating path $(A^\beta \ldots)$ issuing from an unsaturated point A with an initial β-edge, then the new colouring will be again a suitable colouring of the graph. That is, we have to prove that if there is in the new colouring a singular path, then there is such a path even in the original colouring. Let us suppose, therefore, the existence of a path $(B^\beta - {}^\beta C)$ in the new colouring, where B and C are both unsaturated points and their deficiency-sum is greater than 1.

In the original colouring the points A, B, C are not only all unsaturated, but any two of them have a deficiency-sum at least two. This is a simple consequence of the fact that in the original colouring A has one greater deficiency, while all other points have the same deficiency as in the new colouring.

It is sufficient to prove that in the old colouring some two of A, B, C are connected by a $\beta\beta$-path, i. e., a singular path exists.

Let us consider a point D on the path $(A^\beta \ldots)$ such that in $(A^\beta \ldots) = (A^\beta - {}^u D) + (D^u \ldots)$ we have $u = \beta$ and the infinite path $(D^{\bar{u}} \ldots)$ contains no point of $(B^\beta - {}^\beta C)$. The existence of such a D is implied by the finite degree of the graph.

We consider only the points and edges of the alternating path $(A^\beta - {}^\beta D)$, taken in the original colouring, and leave all other points and edges of the graph aside. Then the number of β-edges issuing from A and D exceeds with one

the number of α-edges issuing from A and D, while for all other points the two corresponding numbers are equal. The same is true for the path $(B^\beta - {}^\beta C)$ in the new colouring with respect to B and C.

Let us consider the points and edges of the paths $(A^\beta - {}^\beta D)$ and $(B^\beta - {}^\beta C)$ in the original resp. in the new colouring, twice considering the common edges. The coinciding edges of these two paths have different colours in the two paths, while all other edges, even those of the path $(B^\beta - {}^\beta C)$, have the original colouring. It is evident that by omitting the coinciding pairs of edges the difference of the numbers of the α- and β-edges incident to any point does not chan e. The finite subgraph Φ arising by this omission out of the paths $(A^\beta - {}^\beta D)$ and $(B^\beta - {}^\beta C)$ has the property that to all of its points the number of inciden β edge and α edge are equal except A, B, C, D to which one more (if 2, 3 or 4 of them are coinciding, then 2, 3 or 4 more) β-edge belongs. We shall show that in the graph Φ (whose edges have their original colours) there exists a $\beta\beta$-path which is a singular path of G.

From what has been said about the graph Φ we infer that we can start from A along an alternating path with an initial β-edge and we must arrive at a point whence we cannot proceed so that the path shall be alternating. This point is surely one of B, C, D. If this point is either B or C, then we are ready. If this point is D, then omit the edges of the path from A to D just constructed. The remaining subgraph of Φ has the property that to all of its points the same number of β-edges as α-edges is incident with the exception of the points B and C to which one more (if $B = C$, two more) β-edge is incident. By the process used above we get a $\beta\beta$-path connecting B and C which is a singular path of G. This completes the proof.

(Received 16 May 1950)

О ФАКТОРИЗАЦИИ „ГРАФОВ"

Т. ГАЛЛАИ (Будапешт)

(Резюме)

"Графом" называется некоторое множество точек и ребер, причем к каждому ребру принадлежат две точки. (Streckenkomplex, simplicial 1-complex.) Валентным числом подграфа называется число тех ребер, только одна точка которых принадлежит к подграфу. Число упирающихся в одну точку ребер называется степенью точки. Граф называется правильным, если степени всех его точек равны между собою. Общая степень точек правильного графа называется степенью графа. Фактором графа называется правильный подграф содержащий все точки графа.

Важной проблемой теории графов является вопрос существования факторов. В этом направлении известны исследования *Петерсена*, *Кенига*, *Беблера*, *Холля-Радо* и *Туттэ*. Теоремы этих авторов обладают различным характером в зависимости от того, является ли степень графа и фактора четной, или нечетной.

В настоящей работе вышеупомянутые теоремы доказываются, исходя из общего основания и дается новое условие существования факторов, существенно независимое от четного характера степени. Теоремы доказываются сначала для конечных графов, а затем мы их обобщаем и на бесконечные графы конечной степени (§). В наших утверждениях важную роль играет минимум зацепных чисел, принадлежащих подграфам состоящим из нечетного числа точек (ξ).

После вступительных теорем первых четырех параграфов мы доказываем в пятом параграфе теоремы *Тутта, Холля* и *Радо* относящиеся к существованию фактора первой степени.

В шестом параграфе мы доказываем следующую теорему:

ТЕОРЕМА II. *Конечный, связанный, правильный граф степени $\nu = \sigma + \tau$ обладает фактором степени σ и τ если $\xi \geq \max\left(\frac{\nu}{\sigma}, \frac{\nu}{\tau}\right)$ предполагая, что при σ и τ — нечетных, граф имеет четное число точек.*

В седьмом параграфе, принимая во внимание четность или нечетность ν, σ, τ, мы доказываем следующие следствия предыдущей теоремы, которые относятся к конечным, связанным, правильным графам:

Если степень ν графа нечетная, то существует для каждой четной σ, удовлетворяющей условию: $\sigma \leq \nu - \frac{\nu}{\xi}$ фактор степени σ.

Если степень ν графа нечетная и граф не имеет моста $(\xi > 1)$ то существует фактор четной степени не превосходящий $\frac{2}{3}\nu$ и фактор нечетной степени не меньше $\frac{\nu}{3}$.

Если граф имеет четное число точек, его степень ν четная и у него нет двойного моста $(\xi > 2)$, то существует для каждой σ, удовлетворяющей требованию: $\frac{\nu}{4} \leq \sigma \leq \frac{3\nu}{4}$ фактор степени σ.

Akadémiai Kiadó (Budapest VI, Sztálin-út 31) Felelős: Mestyán János

53 161 — Egyetemi Nyomda, Budapest. (F.: Mező Sándor)

Les Acta Mathematica paraissent en russe, français, anglais et allemand et publient des mémoires du domaine des sciences mathématiques.

Les Acta Mathematica sont publiés, à époques indéterminées, en cahiers qui seront réunis en volumes de 300 à 500 pages. Il paraît, en général, un volume par an.

Les manuscrits, autant que possible écrits à la machine, doivent être envoyés à l'adresse suivante:

Acta Mathematica, Budapest 62, Postafiók 440.

Toute correspondance doit être envoyée à cette même adresse.

Le prix de l'abonnement est 60 forints par an.

On peut s'abonner à l'entreprise de commerce extérieur des livres et journaux „Kultúra" (Budapest V, Akadémia-u. 10. Compte courant No. 929040) ou à l'étranger chez tous les représentants ou dépositaires.

The Acta Mathematica publish papers on mathematics in Russian, French, English and German.

The Acta Mathematica appear in parts of various size, making up volumes of 300—500 pages. On the average, one volume is published annually.

Manuscripts should, if possible, be typewritten and addressed to

Acta Mathematica, Budapest 62, Postafiók 440.

Correspondence with the editors or publishers should be sent to the same address.

The rate of subscription to the Acta Mathematica, is 60 forints a volume. Orders may be placed with „Kultúra" Foreign Trade Company for Books and Newspapers (Budapest V, Akadémia-u. 10. Account No. 929040) or with representatives abroad.

Die Acta Mathematica veröffentlichen Abhandlungen aus dem Bereiche der mathematischen Wissenschaften in russischer, französischer, englischer und deutscher Sprache.

Die Acta Mathematica erscheinen zwanglos in Heften. Mehrere Hefte bilden einen Band von 20—30 Bogen. Im allgemeinen erscheint jährlich ein Band.

Die zur Veröffentlichung bestimmten Manuskripte sind, möglichst mit Maschine geschrieben, an folgende Adresse zu senden:

Acta Mathematica, Budapest 62, Postafiók 440.

An die gleiche Anschrift ist auch jede für die Redaktion und den Verlag bestimmte Korrespondenz zu senden.

Abonnementspreis pro Band 60 Forints. Bestellbar bei dem Buch- und Zeitungs-Aussenhandels-Unternehmen „Kultúra" (Budapest V, Akadémia-u. 10. Bankkonto Nr.: 929040) oder bei seinen Auslandsvertretungen und Kommissionären.

26·— Ft

INDEX